

	1	921	68.5	960	28	CC344833	CG1A020TV
	2	839.4	62.5	923	29	CG240397	OGYCI27TV
	3	762.2	56.0	818	29	CG453578	CG9AD40TV
	4	752	56.0	873	28	CC344824	CG1A020TH
						CC344833	OG1A020TV

```
FEATURES
  source
    1..960
      /location=Qualifiers
      /organism="Zea mays"
      /mol_type="genomic DNA"
      /strain="B73"
      /db_xref="taxon:4577"
      /clone="ZMBMa0359C15"
      /clone_lib="ZM 0.7 1.5 kb"
      /note="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
      methylation filtered genomic DNA library"

ORIGIN
  Query Match
  Best Local Similarity 68.5%; Score 921; DB 28; Length 960;
  Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 35 GTCCCTTCCCATACATGTCGACCTCCACCGCGGAGCACCAGTCGCGGCGCAC 94
  Db 1 GTCCCTTCCCATACATGTCGACCTCCACCGCGGAGCACCAGTCGCGGCGCAC 60
  QY 95 GCGCTCCGCGAGAGCTGGCGCGCTCATCGAGGCTCGGCTCTCTCTACAGCGC 154
  Db 61 GCGCTCCGCGAGAGCTGGCGCGCTCATCGAGGCTCGGCTCTCTCTACAGCGC 120
  QY 155 TCCAGCGCGGACCTGGGAGGACAGAGTGCGCTTCTATGAGGCGTTCTCGGCCACG 214
  Db 121 TCCAGCGCGGACCTGGGAGGACAGAGTGCGCTTCTATGAGGCGTTCTCGGCCACG 180
  QY 215 CGCGCTCCGCGCGCATCCGAGACACCTCTCTCCCGGTTTCCAGCGACGACTCC 274
  Db 181 CGCGCTCCGCGCGCATCCGAGACACCTCTCTCCCGGTTTCCAGCGACGACTCC 240
  QY 275 TCCCGACGAGCGCAGCGCGGAGCGCAGCTCTCAGCTCGCTCTCGAGACCTCTCG 334
  Db 241 TCCCGACGAGCGCAGCGCGGAGCGCAGCTCTCAGCTCGCTCTCGAGACCTCTCG 300
  QY 335 CGGGTTTCAGCGCGCTCGTCGAGACATCAAGATCGCGCCATCATCGTGGCCACCGA 394
  Db 301 CGGGTTTCAGCGCGCTCGTCGAGACATCAAGATCGCGCCATCATCGTGGCCACCGA 360
  QY 395 GTTCGCGGAGCGCTTACATCGCGAGTGTCTCGCAAGACCGCGGAGCACCAGCGTTC 454
  Db 361 GTTCGCGGAGCGCTTACATCGCGAGTGTCTCGCAAGACCGCGGAGCACCAGCGTTC 420
  QY 455 TGCTCGGATTCGCGCTCTCGCGCTCCGAGTGTCTCGCGCGCGGAGCGCGTGTGCGGA 514
  Db 421 TGCTCGGATTCGCGCTCTCGCGCTCCGAGTGTCTCGCGCGCGGAGCGCGTGTGCGGA 480
  QY 515 CGAGCGCGCGGAGGTGAAGGCATGGAACACCGCGCGCTCGCGCGCTCTCGCGCGCT 574
  Db 481 CGAGCGCGCGGAGGTGAAGGCATGGAACACCGCGCGCTCGCGCGCTCTCGCGCGCT 540
  QY 575 AGTGTATCTCTGCGGAGGAGGATGAGTGTCTGCGCTCGCGCGCGGAGTGTGAGGAG 634
  Db 541 AGTGTATCTCTGCGGAGGAGGATGAGTGTCTGCGCTCGCGCGCGGAGTGTGAGGAG 600
  QY 635 GAAAGGTGAGTCTTGTTCACAGCTGCGGAGCTCAAGCGTGGTTCGAGGAGCAGATC 694
  Db 601 GAAAGGTGAGTCTTGTTCACAGCTGCGGAGCTCAAGCGTGGTTCGAGGAGCAGATC 660
  QY 695 TGTTCCATCTTCTACTCGGCGTCAATTTCTTGGGCTATGATCTGCTGAGTTCGAGCAG 754
  Db 661 TGTTCCATCTTCTACTCGGCGTCAATTTCTTGGGCTATGATCTGCTGAGTTCGAGCAG 720
  QY 755 GCGGAGATGGGGTGGGTTGAGGTGAGCTGAGTGTGAGCTTTCGCCATGTGGCGGAGGTG 814
  Db 721 GCGGAGATGGGGTGGGTTGAGGTGAGCTGAGTGTGAGCTTTCGCCATGTGGCGGAGGTG 780
  QY 815 ATGGGAGTGTGACACACACTTCTCGGCGGCTCTGCTCGTGTGATCAAGTTCGTTCTG 874
  Db 781 ATGGGAGTGTGACACACACTTCTCGGCGGCTCTGCTCGTGTGATCAAGTTCGTTCTG 840
  QY 875 ACATTGTTCCGAGAGCTCTCTCATAGCGACCTTTGGGTCTCTTCTTAAGAGAGGATCTGG 934

RESULTS 2
CG240397
LOCUS
DEFINITION
  CG240397 ZM 0.7 1.5 kb Zea mays genomic clone ZMBMa0789E05,
  genomic survey sequence.
ACCESSION
  CG240397
VERSION
  CG240397.1 GI:34140283
SOURCE
  GSS.
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
  1 (bases 1 to 923)
  White, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
  Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Other_GSSs: OGYC127TH
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@cigr.org
  Seq primer: TP
  Class: sheared ends.
  Location/Qualifiers
    1..923
      /organism="Zea mays"
      /mol_type="genomic DNA"
      /strain="B73"
      /db_xref="taxon:4577"
      /clone="ZMBMa0789E05"
      /clone_lib="ZM 0.7 1.5 kb"
      /note="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
      methylation filtered genomic DNA library"

ORIGIN
  Query Match
  Best Local Similarity 62.5%; Score 839.4; DB 29; Length 923;
  Matches 875; Conservative 0; Mismatches 21; Indels 3; Gaps 2;
  QY 42 TCCCATACCATGTCGACCTCCACCGCGGAGACCAAGTCGCGGCGCTCCAGCCTC 101
  Db 24 TCCCATACCATGTCGACCTCCACCGCGGAGACCAAGTCGCGGCGCTCCAGCCTC 83
  QY 102 CGCAGCAAGTCGCGCGCTCATCGAGGCTCGCGCTCTTCTACAGCGCTCCAGGC 161
  Db 84 CGCAGCAAGTCGCGCGCTCATCGAGGCTCGCGCTCTTCTACAGCGCTCCAGGC 143
  QY 162 CGGCGACCGTGGGAGCAGAGTTCGCTTCTATGAGCGTTCCTCCGCCACGCGCGT 221
  Db 144 CGGCGACCGTGGGAGCAGAGTTCGCTTCTATGAGCGTTCCTCCGCCACGCGCGT 203
  QY 222 CCGCGCGCGATCCGAGACACTTCTTCCCGGCTTCACGCGAGCTCTCTCCCGAC 281
  Db 204 CCGCGCGCGATCCGAGACACTTCTTCCCGGCTTCACGCGAGCTCTCTCCCGAC 263
  QY 282 CGAGCGCAGCGCGGAGCGGATCTCTCACTCGTCTCGAGCCTCTCGCGGGGT 341
  Db 264 CGAGCGCAGCGCGGAGCGGATCTCTCACTCGTCTCGAGCCTCTCTCGCGGGGT 323
  QY 342 TCAGCGCGCTCGCTCGCAGACATCAAGATCGCGGCGCATCAGTGGCGCACCGAGTTCGC 401
```





```

ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..873
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM0359C15"
/clone_lib="ZM 0.7 1.5 kb"
/notes="Vector: pBCSK-; Site 1: HindII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 56.0%; Score 752; DB 28; Length 873;
Best Local Similarity 100.0%; Pred. No. 1.1e-105; Indels 0; Gaps 0;
Matches 752; Conservative 0; Mismatches 0;
QY 204 CTCGCCACGCGCGCGTCCGCGCGCATCCGAGACACCTTCTCCCGGTTCCACGG 263
DB 873 CTCGCCACGCGCGCGTCCGCGCGCATCCGAGACACCTTCTCCCGGTTCCACGG 814
QY 264 CAGCGACTCTCCACGAGCGCAGCGCGGAGCGCATCTCTCTCTCTCTCTCTCTCT 323
DB 813 CAGCGACTCTCCACGAGCGCAGCGCGGAGCGCATCTCTCTCTCTCTCTCTCTCT 754
QY 324 CGACTCTCTCGCGGCTTCAGCGCGCTCGCGCGCATCAAGATCGGCGCCATCAC 383
DB 753 CGACTCTCTCGCGGCTTCAGCGCGCTCGCGCGCATCAAGATCGGCGCCATCAC 594
QY 384 GTGGCCACCGAGTCCGCGGAGCCCTACATCGCCAAAGTGCCTCGCCAAAGACCGCGGAC 443
DB 693 GTGGCCACCGAGTCCGCGGAGCCCTACATCGCCAAAGTGCCTCGCCAAAGACCGCGGAC 634
QY 444 CAGCGGCTCTCTCGGATTCGCGCTTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCG 503
DB 633 CAGCGGCTCTCTCGGATTCGCGCTTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCG 574
QY 504 CGTGTCCGCGACGCGCGCGCGGATGAAGGCATGACGACACCGCGCGCTCGCGCGCT 563
DB 573 CGTGTCCGCGACGCGCGCGGATGAAGGCATGACGACACCGCGCGCTCGCGCGCT 514
QY 564 GCTCCGCGCTACGTGTATTCGTTGCCGACGAGGGATGGACTGTGCGCTCGCGCGCG 623
DB 513 GCTCCGCGCTACGTGTATTCGTTGCCGACGAGGGATGGACTGTGCGCTCGCGCGCG 454
QY 624 GGTGTACGGAGAAAGTGGAGTCTTGTACAGCTCGCGGCTCAAGCGCTGGTTCGA 583
DB 453 GGTGTACGGAGAAAGTGGAGTCTTGTACAGCTCGCGGCTCAAGCGCTGGTTCGA 394
QY 684 GGAGCAGACTCTGTTCACACTTCTACTCGCGGTGCAATCTCTCTGGGTATGATGCTGCTGC 743
DB 393 GGAGCAGACTCTGTTCACACTTCTACTCGCGGTGCAATCTCTCTGGGTATGATGCTGCTGC 334
QY 744 AGTCGACGACGCGGAGATGGGGTGGGTTGACGGTGAAGCTGGTGGACTTTGCCCATGT 803
DB 333 AGTCGACGACGCGGAGATGGGGTGGGTTGACGGTGAAGCTGGTGGACTTTGCCCATGT 274

```

---

```

QY 804 GCGCCAGGGTGATGGGTGATTGACACAACTTCTCTGGCGGGCTCTGCTCGTGATCAA 863
DB 273 GCGCCAGGGTGATGGGTGATTGACACAACTTCTCTGGCGGGCTCTGCTCGTGATCAA 214
QY 864 GTTCGTTTCTGACATTTGTCGGAGACTCTCATACGACGCTTTGGGTCTCTTTAAGA 923
DB 213 GTTCGTTTCTGACATTTGTCGGAGACTCTCATACGACGCTTTGGGTCTCTTTAAGA 154
QY 924 GAGGATCTCTGGCATTTTCGATTTGATAACAAAG 955
DB 153 GAGGATCTCTGGCATTTTCGATTTGATAACAAAG 122
RESULT 5
CC724960 867 bp DNA linear GSS 23-JUN-2003
LOCUS OGLAB231V ZM 0.7 1.5 kb Zea mays genomic clone ZM0307C22,
DEFINITION genomic survey sequence.
ACCESSION CC724960 GI:32143893
VERSION CC724960.1
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..867
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM0307C22"
/clone_lib="ZM 0.7 1.5 kb"
/notes="Vector: pBCSK-; Site 1: HindII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 48.0%; Score 644.8; DB 29; Length 867;
Best Local Similarity 89.4%; Pred. No. 3.1e-89;
Matches 753; Conservative 0; Mismatches 27; Indels 62; Gaps 3;
QY 2 CACGAGGTGAGTCCGTCACCCCTCGCGCCATAGTCCCTTCCCATACCATGTCGACC 61
DB 87 CACCATGCGCACCCGTCACCCCTTCTGCTCCATAG-----TCCCCATACCATGTCGACC 140
QY 62 TCACCCCGCGGAGACCAAGTCCGCGGCCACCGCGCTCCCGCAGCAGCTGGGCCCGC 121
DB 141 TCACCCCGCGGAGACCAAGTCCGCGGCCCTCCCGCAGCAGCTGGGCCCGC 200
QY 122 TCATCGAGCGTCCGCGCTCTTCTAAGCCGCTCCAGCGCGCGACCGTGGGAGCAG 181
DB 201 TCATCGAGCGCTCTGCGCTCTTCTAAGCCGCTCCAGCGCGCGACCGTGGGAGCAG 260
QY 182 AGGTGCGCTTCTATGAGGCGTTCTCCGCGCCACCGCGCGTCCCGGCCCGCATCCGAGACA 241
DB 261 AGGTGCGCTTCTATGAGGCGTTCTCCGCGCCACCGCGCGTCCCGGCCCGCATCCGAGACA 320

```



```
Db      676 TGGTCCACCTTTACTCGGGGGCGATTCTTCTGGGCTATGAAGCTAGTGCCGTTGCAACA 735
QY      755 GCGG-AGATGGGGTGGGGTGAAGTGAAGCTGGGTGGACTTTGCCCATGTGGCCGAGGGT 813
Db      736 CCGGAAGGTAAAGGGTGGGGTGAAGGTAAACATGGGGGCCCTTTGCCATGTGGCCCAAGG 795
QY      814 GATGGG 819
Db      796 GAAGGG 801

RESULT 7
LOCUS   CG284572
DEFINITION 719 bp DNA linear GSS 25-AUG-2003
          OGWI61TH ZM_0.7_1.5_KB Zea mays genomic clone ZMBEWA0584K02,
          genomic survey sequence.
ACCESSION  CG284572
VERSION    CG284572.1 GI:34198786
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
           clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 719)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
           Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
           Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
           Consortium for Maize Genomics
           Unpublished (2002)
TITLE     Other GSSs: OGWI61TV
JOURNAL   Contact: Cathy Whitelaw
COMMENT   TIGR
           9712 Medical Center Drive, Rockville, MD 20850, USA
           Tel: 301-838-5843
           Fax: 301-838-0208
           Email: whitelaw@tigr.org
           Seq primer: TR
           Class: sheared ends.

FEATURES             source
     source
     1..719
     /organism="Zea mays"
     /mol_type="genomic DNA"
     /strain="B73"
     /db_xref="taxon:4577"
     /clone="ZMBEWA0584K02"
     /clone_lib="ZM_0.7_1.5_KB"
     /note="Vector: pBGSK-; Site_1: HincII; 0.7-1.5 kb
           methylation filtered genomic DNA library"

ORIGIN
Query Match          45.6%; Score 613.2; DB 29; Length 719;
Best Local Similarity 90.4%; Pred. No. 2.3e-84;
Matches 700; Conservative 0; Mismatches 18; Indels 56; Gaps 2;

QY      50 CCATGTCGGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGCGCTCCGCCAGCA 109
Db      1 CCATGCCGACCTCCACCGCGCGGAGCACCAAGTCGCGGTCACCGCGCTCCGCCAGCA 60
QY      110 AGCTGGGCGCGTCAATCGACGGCTCGGCTCTTCTAAGCGCTCAAGCGCGCGGAC 169
Db      61 AGCTGGGCGGACATCATCGACGACTCTGGGCTCTTCTTACAAGCGCTCCAGGCGCGGAC 120
QY      170 GTGGGAGACAGAGTGCCTTCTATGAGGCGTCTTCGCCCCACGCGCGCTCCCGGCC 229
Db      121 GTGGGAGACAGAGTGCCTTCTATGAGCGCTTCTTCGCCCCACGCGCGCTCCCGGCC 180
QY      230 GCATCCGAGACACCTTCTTCCCGCGTTCACGGCACGCGACTCTCTCCCAACCGAGGGCG 289
Db      181 GCATCCGAGACACCTTCTTCCCGCGTTCACGGCACGCGACTCTCTCCCAACCGAGGGCG 240
QY      290 AGCCGGGAGCGGACCTCTCACTCGTCTCTGACGACCTCTCTCGGGGGTTTCAAGGCG 349
```

```
Db      241 AGCCGGGAGCGCATCGGCACCTCTCTCGACGACCTCTCTCGCGGGTTTGAGGCG 300
QY      350 CTTGGTGCAGACATCAAGATCGGCGCCATCACTGTGCCACCGAGTTTCGCGGAGCCCT 409
Db      301 CTTGGTGCAGACATCAAGATCGGTGCCATCACTGTGCCATCACTGTGCCATCACTGT 336
QY      410 ACATCGCCCAAGTGCCTCGCCCAAGGACCGCGGACCCACGAGCGTTCTGTCTCGGATTCGCGG 469
Db      337 -----GACCACGAGCGTTCTGTCTCGGATTCGCGG 365
QY      470 TCTCGGCGTCCGAGTCTCGGCCCCGAGGCGCGCTGTGGCGACCGAGCGCCCGGAGG 529
Db      366 TCTCGGCGTCCGAGTCTCGGCCCCGAGGCGCGCTGTGGCGACCGAGCGCCCGGAGG 425
QY      530 TGAAGCCATCGACACCGCGCGCTCGCGCGCTCTCCGCGCTACCTGTCTCATCCG-TT 588
Db      426 TGAAGCTATGACATTGTCTCGGCGTCTCGCGCGCTCTCCGCGCTACCTGTCTCATCCGTT 485
QY      589 GCCGACGAGGGGATGGACTGTGCGCTCGCGCGCGGCTGTACGGAGGAAAAGGTGGAGTC 648
Db      486 GCCGACGAGGGGATGGACTGTGCGCTCGCGCGCGGCTGTACGGAGGAAAAGGTGGAGTC 545
QY      649 TTGTCACAGTCTCGCGAGCTCAAGCGTGGTTCGAGGAGCAGACTCTGTTCACCTTCTAC 708
Db      546 TTGTCACAGTCTCGCGAGCTCAAGCGTGGTTCGAGGAGCAGACTCTGTTCACCTTCTAC 605
QY      709 TCGGCGTCACTTCTCTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 768
Db      606 TCGGCGTCACTTCTCTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
QY      769 GGGGTACGCTGAAGCTGTGGACTTTTCCCATGTGGCCGAGGGTGTATGGGGTG 822
Db      666 GGGGTAAACAGTGAAGCTGTGGACTTTTCCCATGTGGCCGAGGGTGTATGGGGTG 719

RESULT 8
LOCUS   CG776236
DEFINITION 591 bp DNA linear GSS 29-OCT-2003
          1123005E08.x1 1123 - RescueMu Grid L Zea mays genomic, genomic
          survey sequence.
ACCESSION  CG776236
VERSION    CG776236.1 GI:38034108
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
           clade; Panicoideae; Andropogoneae; Zea.
           1 (bases 1 to 591)
           Walbot,V.
           Maize genomic sequences found using engineered RescueMu transposon
           Unpublished (2001)
           Contact: Walbot V
           Department of Biological Sciences
           Stanford University
           855 California Ave, Palo Alto, CA 94304, USA
           Tel: 650 723 2227
           Fax: 650 725 8221
           Email: walbot@stanford.edu
           Plate: 1123005 column: 3
           Class: transposon-tagged.
           Location/Qualifiers
             1..591
             /organism="Zea mays"
             /mol_type="genomic DNA"
             /cultivar="mixed background W23/A188/B73/X55"
             /db_xref="taxon:4577"
             /dev_stage="leaf"
             /lab_host="DH10B"
             /clone_lib="1123 - RescueMu Grid L"
             /note="Organ: leaf; Vector: RescueMu (engineered from
             pBlueScript backbone); Site_1: BamHI; Site_2: BglIII;
```

RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid 1 was grown in Molokai in 2001. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

```
ORIGIN
Query Match          43.9%; Score 589.4; DB 29; Length 591;
Best Local Similarity 99.8%; Pred. No. 1e-80; Indels 0; Gaps 0;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 25 CGCGCCCATAGTCCCTTCCCTCCCATACCATGTCGAGCTCCACCGCGCGAGCACCAAGTC 84
DB 1 CGCGCCCATAGTCCCTTCCCTCCCATACCATGTCGAGCTCCACCGCGCGAGCACCAAGTC 60

QY 85 GCGGCGCCACCGCGCTCGCGAGCAAGCTGGGCGCGCTCATCGACGGCTCGGGCTTTC 144
DB 61 GCGGCGCCACCGCGCTCGCGAGCAAGCTGGGCGCGCTCATCGACGGCTCGGGCTTTC 120

QY 145 TACAAGCGCGTCCAGGCGCGCGACCGTGGGAGCAGCAGAGTTCGCTTCTATGAGCGTTC 204
DB 121 TACAAGCGCGTCCAGGCGCGCGACCGTGGGAGCAGCAGAGTTCGCTTCTATGAGCGTTC 180

QY 205 TCGCGCCACCGCGCGCTCGCGGCGCGCATCGAGACACCTTCTTCCCGGTTCCAGCGC 264
DB 181 TCGCGCCACCGCGCGCTCGCGGCGCGCATCGAGACACCTTCTTCCCGGTTCCAGCGC 240

QY 265 ACGGGACTCTCCCGCCAGGCGCGACCGCGGAGCGCATCTCATCTGCTCTTCGAC 324
DB 241 ACGGGACTCTCCCGCCAGGCGCGACCGCGGAGCGCATCTCATCTGCTCTTCGAC 300

QY 325 GACCTCTCGCGGGTTTCAGGCGCGCTCGGTGCGAGACATCAAGATCGCGCGCATCAG 384
DB 301 GACCTCTCGCGGGTTTCAGGCGCGCTCGGTGCGAGACATCAAGATCGCGCGCATCAG 360

QY 385 TGGCCACCGAGTTTCGCGGAGCGCTATACATCGCCAAAGTTCCTCGCCAAAGACCGCGGACC 444
DB 361 TGGCCACCGAGTTTCGCGGAGCGCTATACATCGCCAAAGTTCCTCGCCAAAGACCGCGGACC 420

QY 445 ACAGCGGTTCTGTCGAGTTTCGGGTCTCGGCTCGAGTTCGCGCGCGCGGCGGC 504
DB 421 ACAGCGGTTCTGTCGAGTTTCGGGTCTCGGCTCGAGTTCGCGCGCGCGGCGGC 480

QY 505 GTGTGGCGGACGAGCGCGCGAGTGAAGCCATGACACCGCGCGCTCGCGCGGTG 564
DB 481 GTGTGGCGGACGAGCGCGCGAGTGAAGCCATGACACCGCGCGCTCGCGCGGTG 540

QY 565 CTCGCGGCTACGTGTCTATCGTTTCGCGAGCGAGGGATGAGTGTGCGTTC 615
DB 541 CTCGCGGCTACGTGTCTATCGTTTCGCGAGCGAGGGATGAGTGTGCGTTC 591
```

```
RESULT 9
CA245555
LOCUS SCBGLF5081A03.9 Saccharum officinarum FL5 Saccharum officinarum
DEFINITION CNA clone SCBGLF5081A03 5', mRNA sequence.
ACCESSION CA245555
VERSION CA245555.1 GI:35324301
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Saccharum.
1 (bases 1 to 656)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
```

JOURNAL  
COMMENT

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bccc.unesp.br  
Plate: 081 row: A column: 03  
Seq primer: T7 Promoter Primer.

FEATURES  
source

1..656  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCBGLF5081A03"  
/lab\_host="DH10B"  
/clone\_lib="Saccharum officinarum FL5"  
/notes="Organ: Developed inflorescence (20cm-long) without  
rachis; Vector: pSport1; Site 1: SalI; Site 2: NotI; An  
unidirectional cDNA library generated from [developed  
inflorescence (20cm-long) without rachis]. cDNA was  
prepared from polyA+ mRNA using SuperScript plasmid  
System Kit (Invitrogen). The double-strand cDNAs were  
fractionated in a sepharose Cu-2B 40cm-columns and  
fragments sizing between 0.8 and 1.5 Kb were  
directionally cloned into the vector. Details of each  
source of RNA and library construction can be obtained at  
http://sucest.lad.ic.unicamp.br/public"

## ORIGIN

```
Query Match          42.0%; Score 564.6; DB 14; Length 656;
Best Local Similarity 93.7%; Pred. No. 6.5e-77;
Matches 599; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 37 CCCTTCCCATACCATGTCGAGCTCCGACCTCCACCGCGCGAGCACCAAGTCGCGGCGCACGC 96
DB 18 CCTCTCGCGCTCACCATGTCGACCTCCGCGCGCGAGCACCAAGTCGCGGCGCACGC 77

QY 97 GCCTCGCGCAGCAAGTGGGCGCGCTCATCGAGGTCGCGCTCTTCTTACAGCGCGCTC 156
DB 78 GCCTCGCGCAGCAAGTGGGTCGCTCATCGAGGCTCTGGCTCTTCTACAGCGCGCTC 137

QY 157 CAGGCGCGGACCGTGGGAGCAGAGTTCGCTTCTATGAGCGGTTCTCCGCGCGCACGC 216
DB 138 CAGGTCGGGAGCGCGGAGCAGAGTTCGCTTCTACGAGGCGTTCTCCACCGCGCC 197

QY 217 GCGGTCGCGCGCGCATCCGAGACAGCTTCTCCCGGTTCCACGCGACGCGACTCCTC 276
DB 198 GCGGTCGCGCGCGCATCCGAGACAGCTTCTCCCGGTTTCACGCGACGCGACTCCTC 257

QY 277 CCCACCGAGCGCAGCGCGGAGCGCATCTCTCACTCGCTCTCGACGACCTCTCTGGCG 336
DB 258 CCCACCGAGCGCGCGGAGCGCGCATCTCTCACTCGCTCTCGACGACCTCTCTGGCG 317

QY 337 GGGTTTCAGCGCGCTCGGTTCGAGACATCAAGATCGGCGCCATCAAGTGGCGCACCGAGT 396
DB 318 GGGCTGAGGCGCGCTCGGTTCGCGCGCATCAAGATCGGCGCCATCAAGTGGCGCGCGAGC 377

QY 397 TCGCGGAGCGCTTACATCGCAGTCCCTCGCAAGAGCAGCGGAGCGCGAGGTTCTG 456
DB 378 TCGCGGAGCGCTTACATCGCAGTCCCTCGCAAGAGCAGCGGAGCGCGAGGTTCTG 437

QY 457 CTCGGATTCCGCGCTCCGCGGTCCGAGTCCGCGCGCGCGCGCGCGCGCGCGAGC 516
DB 438 CTCGGATTCCGCGCTCCGCGGTCCGCGGTCCGCGGTCCGCGCGCGCGCGCGCGACT 497

QY 517 GAGCGCGCGGAGGTGAAGGCGCATGAGACCGCGCGCGTCCGCGCGGTCTCCGCGCGCTAC 576
DB 498 GAGCGCGCGGAGGTGAAGGCGCATGAGACCGCGCGCGGTCTCCGCGCGGTCTCCGCGCGCTAC 557
```

```
QY 577 GTGTCATCCGTTCGCGACGAGGGATGAGCTGTGCGCTCGCGCGCGCGGTGTACGAGG- 635
Db 558 GTTTCATCCGTTCGCGACGAGGGATGAGCTGTGCGCTCGCGCGCGCGGTGTACGCGCGC 617
QY 636 AAAAGGTGGAGTCTTGTACACAGCTCGCGAGTCAAGGC 674
Db 618 AAAAGGGGAGTCTTGTACAGCTCGCGAGCTTAAGGC 656

RESULT 10
LOCUS CAL130685 722 bp mRNA linear EST 24-SEP-2003
DEFINITION SCCRT1004G05.g RT1 Saccharum officinarum cDNA clone SCCRT1004G05
5', mRNA sequence.
ACCESSION CAL130685
VERSION CAL130685.1 GI:35013914
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.
1 (bases 1 to 722)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 004 row: G column: 05
Seq primer: T7 Promoter Primer.
location/Qualifiers
1. .722
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCRT1004G05"
/lab_host="DH10B"
/clone_lib="RTL"
/note="Organ: Root tips (0.3cm-long) from adult plants;
Vector: pSport1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Root tips
(0.3cm-long) from adult plants]; cDNA was prepared from
polyA+ mRNA using Superscript Plasmid System kit
(invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

ORIGIN
Query Match 41.6%; Score 559; DB 13; Length 722;
Best Local Similarity 90.9%; Pred. No. 4.6e-76;
Matches 639; Conservative 0; Mismatches 60; Indels 4; Gaps 4;
QY 37 CCCCTTCCCAATACCATGTCGAGCTCCACCGCGGAGACCAAGTCGCGGCCACCGC 96
Db 16 CTTCTGCGCCTCACCATGTCGACCT-CGCGCGCGGAGCACCAAGTCGCGGCCACCGC 74
QY 97 GCTCGGCGCAGAGTGGCGCGCTCATCGAGCGCTCGGCGCTTCTTACAGCGGCTC 156
Db 75 GCTTCGCGCAACAAGCTGGGTTCGCTCATCGAGCGCTCTGGCGCTTCTTACAAGCGGCTC 134
QY 157 CAGGCGCGCGACCGTGGGAGCAGCAGGTCGCGCTTCTATGAGCGCTTCTCCGCCACCGC 216
```

```
Db 135 CAGTCCGGGACCCCGGGAGACGAGCTCGCCTTCTACGAGGCGTTCTCCACCACGCC 194
QY 217 GCCCTCCCGGCCCGCATCCGAGACACCTTCTTCCCGCGGTTCACCGCACCGACTCTC 276
Db 195 GCCGTCCCGGCCCGCATCCGGGACACCTTCTTCCCGCGGTTCACGGCACCGACTCTC 254
QY 277 CCACCGAGCGGACCGCGGGAGCGGATCCTCCTCGCTCTCGAGGACTCTCTCGCG 336
Db 255 CCCACCGAGCGCGCGGAGCGGAGCCGATCCGACCTCGCTCTCGAGGACTCTCTCGCG 314
QY 337 GGGTTTCAGGGCGCCCTCGTCCGACAGATCAAGATCGCGGCCATCACTGTCGCCACCGAGT 396
Db 315 GGGTCTGAGGGCGCCCTCGCTCGCGCATCAAGATCGGGCCATCACTGTCGCCCGCGAGC 374
QY 397 TCGCGGAGCGCCCTACATCGCCAAAGTCTCGCCAAAGACCGCGGACCAACGAGGTTCTG 456
Db 375 TCGCGGAGCGCCCTACGTCCGCAAGTCTCGCCAAAGACCGCGGACCAACGAGATTCTG 434
QY 457 CTCGATTCGCGTCTCCGGGTCGAGTCTCGGCCCCCGAGGCGCGCTGTGGCGGACG 516
Db 435 CTCGATTCGCGTCTCCGGGTCGCGGTCTCGGTCCCGAGGGCGCGCTGTGGCGGACT 494
QY 517 GAGCGCCCGGAGGTGAAGGCGCATGACACCGCCCGCGTCCGCGCGCTCTCCGGCGCTAC 576
Db 495 GAGCGCCCGGAGGTGAAGGCGCTGACACCGCGCGCTCCGCGCGCTCTCCGGCGCTTA 554
QY 577 GTGTATCCGTTGCCGACGAGGGATGAGTGTGCGTTCGCGCGCGCGGTGTA-CGAGG 635
Db 555 CGTTTCATCCGTTGCGGACGAGGGATGAGTGTGCGTTCGCGCGCGCGGAGAACCGCGG 614
QY 636 AAAAGGTGGAGTCTTGTTCACAGCTCGCGAGCTCAAGCGGTGTCGAGGAGCAGACTCT 695
Db 615 CAAAAGGGGAGTCTTGTTCACAGCTCGCGGAGCTTAAAGCGTGGGTGAGGA-CAAACTCT 673
QY 696 GTTCCACTTCTACTCGCGGTCGATTTCTTCTGGGCTATGATGCT 738
Db 674 GGATCACTT-TACTCGGGCGGATCTTCTTCGGGCTAATATGCT 715

RESULT 11
AY109355 3374 bp mRNA linear HTC 17-OCT-2002
LOCUS Zea mays CL49_2 mRNA sequence.
DEFINITION AY109355
ACCESSION AY109355.1 GI:21213015
VERSION AY109355.1
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3374)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 3374)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
location/Qualifiers
1. .3374
/organism="Zea mays"
source
```







Db ATCCGGGACACCTTCTTCCCGGGTTTACGGCAGCGAGACTCTCTCCACCGAGGCGGG 121  
QY CCGGGGAGCGGATCCTCACCCTGCTCTCAGCAGCCTCTCTCGGGGTTTACGGGCGCC 351  
Db CCAGGGGAGCGGATCCTCGGACCTCTCTCAGCAGCCTCTCTCGGGGTTTACGGGCGCC 181  
QY TGCGTCCGAGACATCAAGATCGGCGGCATCACTGTGCGCACCGAGTTTCGCGGAGCGCCCTAC 411  
Db TGCGTCCGAGACATCAAGATCGGCGGCATCACTGTGCGCACCGAGTTTCGCGGAGCGCCCTAC 241  
QY ATGCGCAAGTGCCTCCGCAAGGACCGCGGACCAAGAGGCTTCTGCTCGGATTCGCGCTC 471  
Db GTGCGCAAGTGCCTCCGCAAGGACCGCGGACCAAGAGGCTTCTGCTCGGATTCGCGCTC 301  
QY TCGGCGTCCGAGTCTCTCGGCGCGGAGCGGCGGCTGTGCGGAGCGCGCGGAGGTTG 531  
Db TCGGCGTCCGAGTCTCTCGGCGCGGAGCGGCGGCTGTGCGGAGCGCGCGGAGGTTG 361  
QY AAGGCCATGGACACCGCGCGGCTCTCGGCGGCTCTCGGCGGCTCTCGGCGGCTCTCGG 591  
Db AAGGCCATGGACACCGCGCGGCTCTCGGCGGCTCTCGGCGGCTCTCGGCGGCTCTCGG 421  
QY GACGAGGAGTGGACGCTGCTGCGGCGGCGGCTGTACGAGGAGAAAGTGGAGTCTTG 651  
Db GACGAGGAGTGGACGCTGCTGCGGCGGCGGCTGTACGAGGAGAAAGTGGAGTCTTG 481  
QY TCACAGCTCGCGGAGCTCAAGCGTGTGCTGCGGAGGAGCAGCTCTCTCTCTCTCTCT 711  
Db TCACAGCTCGCGGAGCTCAAGCGTGTGCTGCGGAGGAGCAGCTCTCTCTCTCTCTCT 541  
QY GCGTGCATCTTCTGCGGCTATGATGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGG 771  
Db GCGTGCATCTTCTGCGGCTATGATGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGG 601  
QY GTGAGCGTGAAGCTGTGAGCTTTGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAG 831  
Db TTGAGCGTGAAGCTGTGAGCTTTGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAG 660  
QY AACTTCTCTGGG 842  
Db AACTTCTCTGGG 671

RESULT 13  
CC724950/c  
LOCUS  
DEFINITION  
OGLAB23TH ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0307C22,  
genomic survey sequence.  
ACCESSION  
CC724950  
VERSION  
CC724950.1 GI:32143883  
KEYWORDS  
GSS.  
SOURCE  
Zea mays  
ORGANISM  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 722)  
Whiteley, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other GSSs: OGLAB23TV  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@cigr.org  
Seq primer: TR  
Class: sheared ends.  
Location/Qualifiers  
1. .722

/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0307C22"  
/clone\_lib="ZM\_0.7\_1.5\_KB"  
/note="Vector: pBSK-; site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

Query Match  
Best Local Similarity 40.2%; Score 540.8; DB 29; Length 722;  
Matches 638; Conservative 89.6%; Pred. No. 2.8e-73;  
Indels 57; Gaps 3;  
Mismatches 17;  
QY 246 CTTCCCGGTTTCCAGCGCAGCGAGCTCTCTCCACCGAGGCGGAGCGGAGCGCGCA 305  
Db 722 CTTCCCGGTTTCCAGCGCAGCGAGCTCTCTCCACCGAGGCGGAGCGGAGCGCGCA 663  
QY 306 TCCTCACCTCGTCTCGACGACCTCTCTCGGGGTTTTCAGGCGGCTCTCGTCGACAT 365  
Db 662 TCCGACCTCTCTCGACGACCTCTCTCGGGGTTTTCAGGCGGCTCTCGTCGACAT 603  
QY 366 CAAGATCGCGGCGCATACGTTGCGGCCACCGAGTTTCGCGGAGCGGCTTACATGCGCAATGCT 425  
Db 602 CAAGATCGGTGCGCATACGTT-----GACCATGAGGCTTCTGCTCGGATTCGCGGCTCCGAGT 583  
QY 426 CGCAGGAGCGCGGCGGAGCAGGCGGTTCTGCTCGGATTCGCGGCTCTCGGCGTCCGAGT 485  
Db 582 -----GACCATGAGGCTTCTGCTCGGATTCGCGGCTCTCGGCGTCCGAGT 538  
QY 486 CGTCGCGCGCGAGGCGGCGGCTGTGCGGAGCAGGAGCGCGCGGAGTGAAGGCGCATGGAAC 545  
Db 537 CGTCGCGCGCGAGGCGGCGGCTGTGCGGAGCAGGAGCGCGCGGAGTGAAGGCTATGACAT 478  
QY 546 CGCGGCTCTCGCGCGGCTCTCGGCGCTACGTCATCGG-----TTGCGGAGCGGAGTGG 604  
Db 477 TGTGCGGCTCTCGCGCGGCTCTCGGCGCTACGTCATCGGCTTTCGCGAGCGGAGTGG 418  
QY 605 ACTGTCGCTCTCGCGCGGCTGTACGAGGAGGAGGAGGAGTGGAGTCTTGTTCACAGCTCCGCG 664  
Db 417 ACTGTCGCTCTCGCGCGGCTGTACGAGGAGGAGGAGGAGTGGAGTCTTGTTCACAGCTCCGCG 358  
QY 665 AGCTCAAGCGCTGTGCTCGAGGAGCAGACTCTGTTTCCACTTCTACTCGGCGTCTGATCTTC 724  
Db 357 AGCTCAAGCGCTGTGCTCGAGGAGCAGACTCTGTTTCCACTTCTACTCGGCGTCTGATCTTC 298  
QY 725 TGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 784  
Db 297 TGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 238  
QY 785 TGGTGGACTTTCCTCCATGTCGCGGAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 844  
Db 237 TGGTGGACTTTCCTCCATGTCGCGGAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 178  
QY 845 GGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 904  
Db 177 GGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 118  
QY 905 CTTTGGGTCCTTCTTAAGAGGAGTCTTGGCA-----TTTGGATTTGATAAAG 955  
Db 117 CTTTGGGTCCTTCTTAAGAGGAGTCTTGGCA-----TTTGGATTTGATAAAG 66

## RESULT 14

BZ774817  
LOCUS  
DEFINITION  
i51a11.b1 WGS-ZmaysF (DH5a methyl filtered) Zea mays genomic clone  
i51a11, genomic survey sequence.  
ACCESSION  
BZ774817  
VERSION  
BZ774817.1 GI:28952260  
KEYWORDS  
GSS.  
SOURCE  
Zea mays  
ORGANISM  
Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 593)

REFERENCE  
AUTHORS

Rabinowitz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N., Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L., Zutavern, T., McCombie, W.R. and Martienssen, R.A.  
Genomic shotgun sequences from Zea mays (methyl-filtered)  
Unpublished (2002)

TITLE  
JOURNAL  
COMMENT

Contact: W. Richard McCombie

Lita Annerberg Hazen Genome Sequencing Center

Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884

Fax: 516 367 8874

Email: mcombie@cshl.org

Plate: i151 row: a column: 11

Seq primer: -21M13UnivFwd

Class: shotgun

High quality sequence stop: 593.

FEATURES  
source

Location/Qualifiers

1..593

/organism="Zea mays"

/mol\_type="genomic DNA"

/cultivar="B73"

/db\_xref="taxon:4577"

/clone="i151all"

/lab\_host="DH5a"

/clone\_lib="WGS-ZmaysF (DH5a methyl filtered)"

/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I;  
The vector was digested with XbaI and one nucleotide was  
added by fill in the recessive 3' end. The genomic DNA  
was nebulized, end repaired, adaptor ligated and size  
fractionated using sephadex. The resulting fragments were  
between 0.8 and 3 kb and were cloned into the vector (.x/y  
reads in M13mp19, .b/g reads in pUC19). The same ligation  
was transformed into DH5a."

## ORIGIN

Query Match 39.9%; Score 535.6; DB 28; Length 593;  
Best Local Similarity 97.9%; Pred. No. 1.8e-72;  
Matches 564; Conservative 0; Mismatches 9; Indels 3; Gaps 2;

QY 42 TCCCATACCATGTCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGCGCTC 101  
DB 20 TCCCATACCATGTCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGCGCTC 79

QY 102 CGCAGCAAGCTGGGCGCGCTCATCGAGCGCTCCGGCTCTTCTACAGCGCTCCAGGC 161  
DB 80 CGCAGCAAGCTGGGCGCGCTCATCGAGCGCTCCGGCTCTTCTACAGCGCTCCAGGC 139

QY 162 CGCAGCAAGCTGGGCGCGCTCATCGAGCGCTCCGGCTCTTCTACAGCGCTCCAGGC 221  
DB 140 CGCAGCAAGCTGGGCGCGCTCATCGAGCGCTCCGGCTCTTCTACAGCGCTCCAGGC 199

QY 222 CCCGCGCGCATTCGAGACACTTCTTCCCGGTTCCAGCGGACCGGACTCTCTCCCGAC 281  
DB 200 CCCGCGCGCATTCGAGACACTTCTTCCCGGTTCCAGCGGACCGGACTCTCTCCCGAC 259

QY 282 CGAGCGGAGCGCGGAGCGGATCTCTACCTCGTCTCTCGAGCACTCTCTCGGGGTT 341  
DB 260 CGAGCGGAGCGCGGAGCGGATCTCTACCTCGTCTCTCGAGCACTCTCTCGGGGTT 319

QY 342 TCAGCGCGCTCGTCGAGACATCAAGATGGGCGCATCAGTGGGCCACCGAGTTCGCC 401  
DB 320 TGAGCGCGCTCGTCGAGACATCAAGATGGGCGCATCAGTGGGCCACCGAGTTCGCC 379

QY 402 GGAGCGCTTACATCGCCAAAGTGTCTCGCAAGACCGCGGGACCAAGAGCTTCTGTCTGG 461  
DB 380 GGAGCGCTTACATCGCCAAAGTGTCTCGCAAGACCGCGGGACCAAGAGCTTCTGTCTGG 439

QY 462 ATTCGCGCTCTCGGCGCTCGAGTGTCTCGGCGCGCGGCGCGGCTCTTCTACAGCGGCTC 521

DB 440 ATTCGCGCTCT--TCGTCGAGTCGTGCGCCCGGAGGCGCGCTGTGCGGACGAGCG 497  
QY 522 CCCGAGGTGAAGCCATGAGACACCGCGCGGTTCGCGCGGTGCTCCGCGGCTACGTGTC 581  
DB 498 CCCGAGGTGAAGCCATGAGACACCGTCGGGTTCGCGCGGTGCTCCGCGGCTACGTGTC 557  
QY 582 ATTCG-TTTCGCGAGGAGGATGACTGTGCGCTCG 616  
DB 558 ATTCGCTTTCGCGAGGAGGATGACTGCGGCTCG 593

## RESULT 15

LOCUS

CA130686

DEFINITION

5', mRNA sequence.

ACCESSION

CA130686

VERSION

CA130686.1

KEYWORDS

SOURCE

ORGANISM

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bccccenter.fcav.unesp.br  
Plate: 004 row: G column: 06  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers

FEATURES  
source

1..744  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCCPR1004G06"  
/lab\_host="DH10B"  
/clone\_lib="RT1"  
/note="Organ: Root tips (0.3cm-long) from adult plants;  
Vector: pSport1; Site 1: SalI; Site 2: NotI; An  
unidirectional cDNA library generated from [Root tips  
(0.3cm-long) from adult plants]. cDNA was prepared from  
polyA+ mRNA using SuperScript Plasmid System Kit  
(Invitrogen). The double-strand cDNAs were fractionated  
in a sepharose CL-2B 40cm-columns and fragments sizing  
between 0.8 and 1.5 Kb were directionally cloned into the  
vector. Details of each source of RNA and library  
construction can be obtained at  
http://sucest.lad.ic.unicamp.br/public"

## ORIGIN

Query Match 39.1%; Score 525.6; DB 13; Length 744;  
Best Local Similarity 88.6%; Pred. No. 6e-71;  
Matches 592; Conservative 0; Mismatches 74; Indels 2; Gaps 2;

QY 37 CCCCTTCCCATACCATGTCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGC 96  
DB 16 CCTCTCGCCCTCACCATGTCGACCT-CGCGCGCGGAGCACCAAGTCGCGGCCACCGC 74

QY 97 GCCTCCGCGGAGCAAGCTGGGCGCGCTCATCGACGCTCCGCGCTCTTCTTACAGCGGCTC 156  
DB 75 GCCTCCGCGGAGCAAGCTGGGTCGCGCTCATCGACGCTCTGGGCTCTTCTTACAGCGGCTC 134



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2004, 23:22:54 ; Search time 5348 Seconds  
(without alignments)  
10892.490 Million cell updates/sec

Title: US-10-042-894A-7

Perfect score: 1344

Sequence: 1 gcacgagtcagtcgctcac.....ataaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_ov.\*

5: gb\_pat.\*

6: gb\_ph.\*

7: gb\_pl.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_sts.\*

11: gb\_sy.\*

12: gb\_un.\*

13: gb\_vt.\*

14: gb\_vt.\*

15: gb\_vt.\*

16: gb\_vt.\*

17: gb\_vt.\*

18: gb\_vt.\*

19: gb\_vt.\*

20: gb\_vt.\*

21: gb\_vt.\*

22: gb\_vt.\*

23: gb\_vt.\*

24: gb\_vt.\*

25: gb\_vt.\*

26: gb\_vt.\*

27: gb\_vt.\*

28: gb\_vt.\*

29: gb\_vt.\*

30: gb\_vt.\*

31: gb\_vt.\*

32: gb\_vt.\*

33: gb\_vt.\*

34: gb\_vt.\*

35: gb\_vt.\*

36: gb\_vt.\*

37: gb\_vt.\*

38: gb\_vt.\*

39: gb\_vt.\*

40: gb\_vt.\*

41: gb\_vt.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1344	100.0	1344	6	AX513570	AX513570 Sequence
2	865.4	64.4	923	6	AX513568	AX513568 Sequence
3	844.8	62.9	1169	6	AX513564	AX513564 Sequence
4	837.4	62.3	923	6	AX513566	AX513566 Sequence
5	740.4	55.1	3416	6	AX513583	AX513583 Sequence
6	686	51.0	899	6	AX513578	AX513578 Sequence
c 7	546.4	40.7	132170	8	AP005749	AP005749 Oryza sat
8	544.8	40.5	1570	8	AK072296	AK072296 Oryza sat
9	470.4	35.0	643	6	AX513580	AX513580 Sequence
10	298.4	22.2	519	6	AX513581	AX513581 Sequence
11	191.6	14.3	353	6	AX513582	AX513582 Sequence
12	184.6	13.7	1195	6	AX513574	AX513574 Sequence
13	181.6	13.5	25054	8	LUS310150	AXJ10150 Linum usi
14	179	13.3	1105	6	AX513572	AX513572 Sequence
15	156.2	11.6	116893	8	AP004772	AP004772 Oryza sat
16	149.8	11.1	1020	6	AX513576	AX513576 Sequence
17	148.6	11.1	1390	8	AY136378	AY136378 Arabidops
18	148.6	11.1	116783	8	ATT211	AL163912 Arabidops
19	148.2	11.0	905	8	BT000196	BT000196 Arabidops
20	147.8	11.0	115040	8	AC124954	AC124954 Medicago
21	146.6	10.9	1243	8	AY087217	AY087217 Arabidops
22	145.8	10.8	129248	2	AC127018	AC127018 Medicago
c 23	145.6	10.8	903	8	ATH243592	AJ243592 Arabidops
24	145.6	10.8	903	8	AY072621	AY072621 Arabidops
25	145.6	10.8	903	8	AY147936	AY147936 Arabidops
26	145.6	10.8	1166	8	AF412073	AF412073 Arabidops
27	145.6	10.8	1170	8	ATH245521	AJ245521 Arabidops
c 28	145.6	10.8	57246	8	AB010069	AB010069 Arabidops
29	145.4	10.8	861	8	ATH404678	AJ404678 Arabidops
30	145.4	10.8	861	8	AY147935	AY147935 Arabidops
31	143.6	10.7	1130	8	AY085862	AY085862 Arabidops
c 32	83.8	6.2	43428	8	AP004981	AP004981 Lotus cor
33	75.2	5.6	110000	2	LMFLCHR32_06	Continuation (7 of
34	75	5.6	110000	2	LMFLCHR36_03	Continuation (4 of
35	74.8	5.6	296500	1	SCO919128	AL93128 Streptomy
36	73.8	5.5	125020	9	AF429315	AF429315 Homo sapi
c 37	73.6	5.5	110000	2	LMFLCHR36_31	Continuation (32 o
38	73.4	5.5	300800	1	SCO939112	AL939112 Streptomy
39	72.4	5.4	113193	1	AF357202	AF357202 Streptomy
40	72.4	5.4	113193	6	AX703543	AX703543 Sequence
c 41	71.4	5.3	135638	1	AF484556	AF484556 Streptomy
42	70.4	5.2	224	8	AY022706	AY022706 Oryza sat
c 43	70.4	5.2	78220	2	AC023212	AC023212 Homo sapi
44	69.6	5.2	1926	6	AR217867	AR217867 Sequence
45	69.6	5.2	1931	6	AR083152	AR083152 Sequence

# ALIGNMENTS

RESULT 1	AX513570	AX513570	1344 bp	DNA	linear	PAT 05-OCT-2002
LOCUS	AX513570	Sequence 7 from Patent WO20059324.				
DEFINITION	AX513570					
ACCESSION	AX513570					
VERSION	AX513570.1	GI:23559670				
KEYWORDS	..					
SOURCE	Zea mays					
ORGANISM	Zea mays					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD					
	clade; Panicoideae; Andropogoneae; Zea.					
REFERENCE	1					
AUTHORS	Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.					
TITLE	Novel inositol polyphosphate kinase genes and uses thereof					

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 02059324-A 7 01-AUG-2002; (US)  
PIONEER HI-BRED INTERNATIONAL, INC. (US)  
FEATURES  
source Location/Qualifiers  
1. .1344  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"  
52. .921  
/notes="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52954.1"  
/db\_xref="GI:23559671"  
/db\_xref="REMTREMBL:CAD52954"  
/translation="MSDLHPPEHQVAGHRASAKGLPLIDSGSLFYKPLQAGDRGHEV  
VAFYEAFSAHAAPKARIRDTFFPRFGTRLLPTAEQPGFHPHVLVDDLLAGFQAPCV  
ADIKIGAITWPPSPPEFYAKLADKRGVYTAKEADKRTVSVLLGFRVSGVRVVGEGAVWRTPEV  
KAMDTGWPRVLRVRYSSVDEQMDCALAAVYGGKGVLSOLREUKANFPEQTLPHF  
YSASILLGYDAAVAAGDGGGVTVKLVDFAHVAEGDGVIDHNLFLGLCLSLIKFVSDI  
VPETPHQLGPS"

ORIGIN

Query Match	100.0%;	Score 1344;	DB 6;	Length 1344;
Best Local Similarity	100.0%;	Pred. No. 4.1e-193;		
Matches 1344;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			
QY	1	GCACGAGGT	CAGTCCGTCACCCCTCGGGCCCATAGTCCCTTCCCATACCATGTCCGAC	60
Db	1	GCACGAGGT	CAGTCCGTCACCCCTCGGGCCCATAGTCCCTTCCCATACCATGTCCGAC	60
QY	61	CTCCACCCGCGGAGACCAAGTCGCGGGCCACCGCGCCCTCCGCCAGCAGCTGGGCGCG	120	
Db	61	CTCCACCCGCGGAGACCAAGTCGCGGGCCACCGCGCCCTCCGCCAGCAGCTGGGCGCG	120	
QY	121	CTCATCGAGGCTCCGGGCTCTTTCTAAAGCCGCTCCAGGCCGCGGACCGTGGGAGCAC	180	
Db	121	CTCATCGAGGCTCCGGGCTCTTTCTAAAGCCGCTCCAGGCCGCGGACCGTGGGAGCAC	180	
QY	181	GAGTGCCTTCTATGAGGGGTTCTCGGCCACCGCGCGCGTCCGGGCCCGCATCCGAGAC	240	
Db	181	GAGTGCCTTCTATGAGGGGTTCTCGGCCACCGCGCGCGTCCGGGCCCGCATCCGAGAC	240	
QY	241	ACCTTTCTTCCCGGTTTCACGCACGCGACTCTCTCCACCGAGGCGGACGCCGGGAG	300	
Db	241	ACCTTTCTTCCCGGTTTCACGCACGCGACTCTCTCCACCGAGGCGGACGCCGGGAG	300	
QY	301	CGCATCTCTACCTCGTCTTCGACGACCTCCTCGCGGGTTTCAGGCGCCCTCGTCGCA	360	
Db	301	CGCATCTCTACCTCGTCTTCGACGACCTCCTCGCGGGTTTCAGGCGCCCTCGTCGCA	360	
QY	361	GACATCAAGATCGGCGGCATCAGTGGCCACCGAGTTCGCGGAGCCCTACATCGCCAG	420	
Db	361	GACATCAAGATCGGCGGCATCAGTGGCCACCGAGTTCGCGGAGCCCTACATCGCCAG	420	
QY	421	TGCCTGCCAAGACCGCGGGACCAAGAGGTTCTGTCGGAATTCGCGGCTCCGCGCTC	480	
Db	421	TGCCTGCCAAGACCGCGGGACCAAGAGGTTCTGTCGGAATTCGCGGCTCCGCGCTC	480	
QY	481	CGAGTGTGTGGCCCCGAGGGCGCGTGTGCGGACCGAGCGCCCGGAGGTGAAGGCCATG	540	
Db	481	CGAGTGTGTGGCCCCGAGGGCGCGTGTGCGGACCGAGCGCCCGGAGGTGAAGGCCATG	540	
QY	541	GACACCGCGCGTCCGCGCGTCTCCGCGGCTACGTGTCTATCGTTGCGGACGAGGGG	600	
Db	541	GACACCGCGCGTCCGCGCGTCTCCGCGGCTACGTGTCTATCGTTGCGGACGAGGGG	600	
QY	601	ATGGACTGTGCGTTCGCGCGCGGTGTACGAGGAAAAAGGTGGAGTCTTGTCAACAGCTG	660	
Db	601	ATGGACTGTGCGTTCGCGCGCGGTGTACGAGGAAAAAGGTGGAGTCTTGTCAACAGCTG	660	
QY	661	CGCGAGCTCAAGCGTGGTTTCGAGGAGCAGACTCTGTTCACCTTCTACTCGGCGTCAATT	720	
Db	661	CGCGAGCTCAAGCGTGGTTTCGAGGAGCAGACTCTGTTCACCTTCTACTCGGCGTCAATT	720	

QY	721	CTTCTGGGCTATGATGCTGCTGCAGTCGACAGCGGAGATGGGGTGACGGTG	780
DB	721	CTTCTGGGCTATGATGCTGCTGCAGTCGACAGCGGAGATGGGGTGACGGTG	780
QY	781	AAGCTGGTGACATTTGCCATGTGGCGAGGGTGATGGGGTGATTGACCAACTTCTG	840
DB	781	AAGCTGGTGACATTTGCCCATGTGGCGAGGGTGATGGGGTGATTGACCAACTTCTG	840
QY	841	GGCGGCTCTGCTCGCTGATCAAGTTTCGTTTCTGACATTTTCGGGAGATCTCTCATACG	900
DB	841	GGCGGCTCTGCTCGCTGATCAAGTTTCGTTTCTGACATTTTCGGGAGATCTCTCATACG	900
QY	901	CAGCCCTTTGGGTGCTTTCTTAAGAGAGGATCTCTGGCATTTTCGATTTGATAAAGCCCTA	960
DB	901	CAGCCCTTTGGGTGCTTTCTTAAGAGAGGATCTCTGGCATTTTCGATTTGATAAAGCCCTA	960
QY	961	CAAGTTTTCGCGAAAAAGACGGCTCCGAGTTGTGCTGGGTGTGGAGATCTGAGACG	1020
DB	961	CAAGTTTTCGCGAAAAAGACGGCTCCGAGTTGTGCTGGGTGTGGAGATCTGAGACG	1020
QY	1021	GTGCTGGGCCCATTTGGTTGCCCTTTGCTTGGCTCGAACAATACGCGCAACCTGCT	1080
DB	1021	GTGCTGGGCCCATTTGGTTGCCCTTTGCTTGGCTCGAACAATACGCGCAACCTGCT	1080
QY	1081	CCNTTTTTCGCAACCCCTTACTTTCGGAAGAAA	1140
DB	1081	CCNTTTTTCGCAACCCCTTACTTTCGGAAGAAA	1140
QY	1141	TTACGTTGGATCTGGTTTGTGCCACTCGGTTCAGAGGTGTGAAGCATGGAGGAGCGTGT	1200
DB	1141	TTACGTTGGATCTGGTTTGTGCCACTCGGTTCAGAGGTGTGAAGCATGGAGGAGCGTGT	1200
QY	1201	TGATCGCGCAACTGTGTGACGTCTTGGGCTGCGCTTTCGATGGGCTTTTGCTGTC	1260
DB	1201	TGATCGCGCAACTGTGTGACGTCTTGGGCTGCGCTTTCGATGGGCTTTTGCTGTC	1260
QY	1261	TGCGATCCGATGTACTGGAGATCGTAGTGAGCGTCTCTACCTCCAAACGAATCCG	1320
DB	1261	TGCGATCCGATGTACTGGAGATCGTAGTGAGCGTCTCTACCTCCAAACGAATCCG	1320
QY	1321	TCGGATAAAAAAAAAAAAAAAAAA	1344
DB	1321	TCGGATAAAAAAAAAAAAAAAAAA	1344

```

RESULT 2
AX513568
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .923
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
53. -.922
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52953.1"
/db_xref="GI:23559669"
/db_xref="REMTREMBL:CAD52953"
CDS

```

/translation="MPDLHPPEHOVAGHRASAKPGPLIDGSLFYKPLQAGDRGEH  
VAFYFAFAHAAPARIDTREFPHGTRLLPTEAOPGEPHPHVLVDLLAGFAPCV  
ADIKIGALTWPSPSPPEPIAKCLAMDRGTTSVLLGFRVSGVRVVPGEAVWRITERPEV  
KAMDTAGVRRVIRRVSVSVADEGMDCAAAAVYGGKGVLSQLRELKAWFEQTLFHF  
YASILJUGDAAAVAAAGGGGVTVKLVDFARVAAGDGVIDHNFILGGLCSLIFKVS  
VPETPQTQLPGPS"

## ORIGIN

```
Query Match      64.4%; Score 865.4; DB 6; Length 923;
Best Local Similarity 98.8%; Pred. No. 5.5e-121;
Matches 870; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 42 TCCCATACCATGTCGGACCTTCCACCCCGGAGCACCAAGTCGCGCCACCGCGCTC 101
Db 43 TCCCATACCATGTCGGACCTTCCACCCCGGAGCACCAAGTCGCGCTTACCGGCTC 102
QY 102 CGCAGCAAGTGGGCCCGCTCATCGACGGTTCGGGCTCTTTACAGACCGCTCCAGGC 161
Db 103 CGCAGCAAGTGGGCCCGCTCATCGACGGTTCGGGCTCTTTACAGACCGCTCCAGGC 162
QY 162 CGGCGACCGTGGGAGCAGAGTGCCTTCTATAGGCGTTCTCGGCCACCGCGCGT 221
Db 163 CGGCGACCGTGGGAGCAGAGTGCCTTCTATAGGCGTTCTCGGCCACCGCGCGT 222
QY 222 CGGCGCCGCTATCGGAGACCTTCTTCCCGGTTCCAGGCGACGCGACTCTCCCCAC 281
Db 223 CGGCGCCGCTATCGGAGACCTTCTTCCCGGTTCCAGGCGACGCGACTCTCCCCAC 282
QY 282 CGAGCGCAGCCCGGGAGCGCATCTTCTTCTCGTCTCGACGACCTCTCTCGGGGTT 341
Db 283 CGAGCGCAGCCCGGGAGCGCATCTTCTTCTCGTCTCGACGACCTCTCTCGGGGTT 342
QY 342 TCAGCGCGCTCGTTCGAGACATCAAGATCGGCGCATCAGTGGGCCACCGAGTTCGCC 401
Db 343 TGAGCGCGCTCGTTCGAGACATCAAGATCGGCGCATCAGTGGGCCACCGAGTTCGCC 402
QY 402 GGAGCCCTTACATCGCAAGTGCCTCGCAAGACCGCGGACCAAGCGTTCCTCGCG 461
Db 403 GGAGCCCTTACATCGCAAGTGCCTCGCCATGACCGCGGACCAAGCGTTCCTCGCG 462
QY 462 ATTCGGCTCTCCGCGTCCGAGTCTCGGCGCCGAGGCGCGCTGTGGCGGACGAGCG 521
Db 463 ATTCGGCTCTCCGCGTCCGAGTCTCGGCGCCGAGGCGCGCTGTGGCGGACGAGCG 522
QY 522 CCCGAGGTGAGGCCATGACACCGCGCGGTTCGCGCGGTCTCGCGCGCTACGTTGC 581
Db 523 CCCGAGGTGAGGCCATGACACCGCGCGGTTCGCGCGGTCTCGCGCGCTACGTTGC 582
QY 582 ATTCGTTGCCGACGAGGCGATGGAGTGTGCGCTCGCGCGCGGTGTACGAGGAAAGG 641
Db 583 ATTCGTTGCCGACGAGGCGATGGAGTGTGCGCTCGCGCGCGGTGTACGAGGAAAGG 642
QY 642 TGGAGTCTTGTACAGTTCGCGAGCTCAAGCGGTGTTTCGAGGAGCAGACTCTGTTCCA 701
Db 643 TGGAGTCTTGTACAGTTCGCGAGCTCAAGCGGTGTTTCGAGGAGCAGACTCTGTTCCA 702
QY 702 CTTTCTACTCGCGTGCATCTTCTTGGGCTATGATGCTGCTCAGTTCGAGCAGCGGAGA 761
Db 703 CTTTCTACTCGCGTGCATCTTCTTGGGCTATGATGCTGCTCAGTTCGAGCAGCGGAGG 762
QY 762 TGGGGGTGGGGTGCAGGTGAGGTGAGTGTGCGCTTGTGCCCATGTGCCGAGGATGAGG 821
Db 763 TGGGGGTGGGGTGCAGGTGAGGTGAGTGTGCGCTTGTGCCCATGTGCCGAGGATGAGG 822
QY 822 GATTGACACAACTCTTCCTGGGCGGCTCTGCTCGCTGATCAAGTTCGTTTCTGACATTGT 881
Db 823 GATTGACACAACTCTTCCTGGGCGGCTCTGCTCGCTGATCAAGTTCGTTTCTGACATTGT 882
QY 882 TCCGAGACTCTCTATACGAGCCTTTGGGTCCTTCTTAAG 922
Db 883 TCCGAGACTCTCTCAGAGCAGCCTTTGGGTCCTTCTTAAG 923
```

## RESULT 3

AX513564 LOCUS 1169 bp DNA linear PAT 05-OCT-2002  
DEFINITION Sequence 1 from Patent WO2059324.  
AX513564 ACCESSION  
AX513564.1 GI:23559664  
KEYWORDS  
SOURCE  
ORGANISM

Zea mays  
Eukaryota

Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

## REFERENCE

Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.

Novel inositol polyphosphate kinase genes and uses thereof

Patent: WO 02059324-A 1 01-AUG-2002;

PIONEER HI-BRED INTERNATIONAL, INC. (US)

## FEATURES

Location/Qualifiers

1..1169

/organism="Zea mays"

/mol\_type="unassigned DNA"

/db\_xref="taxon:4577"

84..806

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD52951.1"

/db\_xref="GI:23559665"

/translation="MPDLHPPEHOVAGHRASAKPLIDGSLFYKPLQAGDRGEH

VAFYFAFAHAAPARIDTREFPHGTRLLPTEAOPGEPHPHVLVDLLAGFAPCV

ADIKIGALTWPSPSPPEPIAKLXDRGTTSVLLGFRVLRPSRRPRRRVADGAPGGE

GYHRRPPAPALRVIRLTGRMTAKSRRRCTEERKVESCHSASSHRGWSRLCSTS

TRRRFFWAMLLSQAEVGVG"

## ORIGIN

```
Query Match      62.9%; Score 844.8; DB 6; Length 1169;
Best Local Similarity 97.2%; Pred. No. 6.8e-118;
Matches 890; Conservative 0; Mismatches 22; Indels 4; Gaps 3;

QY 42 TCCCATACCATGTCGGACCTTCCACCCCGGAGCACCAAGTCGCGCCACCGCGCTC 101
Db 74 TCCCATACCATGTCGGACCTTCCACCCCGGAGCACCAAGTCGCGCTTACCGGCTC 133
QY 102 CGCAGCAAGTGGGCCCGCTCATCGACGGTTCGGGCTCTTTACAGCGCTCCAGGC 161
Db 134 CGCAGCAAGTGGGCCCGCTCATCGACGGTTCGGGCTCTTTACAGCGCTCCAGGC 193
QY 162 CGGCGACCGTGGGAGCAGAGTTCGCTTCTATAGGCGTTCGCGCCACCGCGCGT 221
Db 194 CGGCGACCGTGGGAGCAGAGTTCGCTTCTATAGGCGTTCGCGCCACCGCGCGT 253
QY 222 CGGCGCCGCTATCGGAGACCTTCTTCCCGCGGTTCACGGCACGCGACTCTCTCCAC 281
Db 254 CGGCGCCGCTATCGGAGACCTTCTTCCCGCGGTTCACGGCACGCGACTCTCTCCAC 313
QY 282 CGAGCGCAGCCCGGGAGCGCATCTCAGTCTGCTCCTCGAAGCCTCTCTCGCGGGTT 341
Db 314 CGAGCGCAGCCCGGGAGCGCATCTCAGTCTGCTCCTCGACACCTCTCTCGCGGGTT 373
QY 342 TCAGGCGCCCTGGTTCGAGACATCAAGATCGGCGCATCAAGTTCGGGCGCTTCCGCG 401
Db 374 TGAGGCGCCCTGGTTCGAGACATCAAGATCGGCGCATCAAGTTCGGGCGCTTCCGCG 433
QY 402 GGAGCCCTTACATCGCCAAAGTGCCTTCGCCAAGGACCGCGGACCAAGCGCTTCTGCTCG 461
Db 434 GGAGCCCTTACATCGCCAAAGTGCCTTCGCCAAGGACCGCGGACCAAGCGCTTCTGCTCG 493
QY 462 ATTTCGGGCTCTCGGCGTTCGAGTTCGTCGGCCCGGAGGGCGCGTGTGGCGGACGAG 521
Db 494 ATTTCGGGCTCT--TGGGTCGAGTTCGTCGGCCCGGAGGGCGCGTGTGGCGGACGAG 551
QY 522 CCCGAGGTCGAGGCGCATGACACCGCGCGCTTCGCGCGGTGCTCTCGGCGCTACGTTGC 581
```

```
Db 552 CCGGAGGTGAAGGCTATGGACACCGTGGCGTCCGCGCGTGTCTCCGCGGCTACGTGTC 611
QY 582 ATCCG-TTCCGACAGAGGGATGGACTGTGCGCTCGCGCGCGGTGTACGAGGAAAG 640
Db 612 ATCCGCTTCCGACAGAGGGATGGACTGCGCGCTCGCGCGCGGTGTACGAGGAAAG 671
QY 641 GTGGAGTCTTGTCACAGCTGCCGAGCTCAAGCGGTGTTCAGAGAGCAGACTCTGTTCC 700
Db 672 GTGGAGTCTTGTCACAGCTGCGAGCTCAAGGCTATGGTGGAGGAGCAGACTCTGTTCC 731
QY 701 ACTTCTACTCGCGTCTGATTTCTTCTGGGCTATGATGCTGCTGAGTGGAGTGGAGGAGCAGACTCTGTTCC 760
Db 732 ACTTCTACTCGCGTCTGATTTCTTCTGGGCTATGATGCTGCTGAGTGGAGTGGAGGAGCAGACTCTGTTCC 791
QY 761 ATGGGGTGGGTGAGCGGTGAAGCTGTGGATTTGTCCTGATGCTGCTGAGTGGAGTGGAGGAGCAGACTCTGTTCC 820
Db 792 GTGGGGTGGGTAAACAGTGAAGTGGTGGATTTGTCCTGATGCTGCTGAGTGGAGTGGAGGAGCAGACTCTGTTCC 851
QY 821 TGATTGACACAACTTCTTCTGGGCGGCTGCTGCTGATCAAGTTCGTTCTGACATTG 880
Db 852 TGATTGACACAACTTCTTCTGGGCGAGCTCTGCTGATCAAGTTCGTTCTGACATTG 911
QY 881 TTCCGAGACTCTCTATAGCGAGCTTTGGGTCCTTTTAAAGAGGATCTTGAATTT 939
Db 912 TTCCGAGACTCTTATAGCGAGCTTTGGGTCCTTTTAAAGAGGATCTTGAATTT 971
QY 940 CGATTGATAACAAAG 955
Db 972 TGATTGATAACAAAG 987

RESULT 4
AXS13566
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
AUTHORS
Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.
TITLE
Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL
Patent: WO 02059324-A 3 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
source
1..923
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
53..736
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD52952.1"
/db_xref="GI:23559667"
/db_xref="translation:MPDLHPPEHVAGHRASAKPGLIDGSLFKYPLQAGDRGEHE
VAFYFAAHAVPARIOTETPRFHGTRLLPTEAOPGPHPLVLDLQAGFEAPCV
ADIKGATWFPSSPEPYAKCLAMDRTSVLLGPRVSGVVRVPEGVWTERPEV
KMDTGVRLKRVLYSSACRRDGLRLARGGGVRRKRSLVTAARAGQVVRGAASVPL
LLGVDSSSL"
ORIGIN
Query Match 62.3%; Score 837.4; DB 6; Length 923;
Best Local Similarity 97.5%; Pred. No. 9.1e-117;
Matches 859; Conservative 0; Mismatches 21; Indels 1; Gaps 1;
QY 42 TCCCATACCATCTCCGACCTCCACCGCGGAGCACCAGTCGCGGCGACCGCGCTC 101
Db 43 TCCCATACCATCTCCGACCTCCACCGCGGAGCACCAGTCGCGGCGACCGCGCTC 102
```

```
QY 102 GGCAGCAAGCTGGGCCCGCTCATCGACGGCTCCGGCCTCTTCTACAAGCGCTCCAGGC 161
Db 103 GGCAGCAAGCTGGGCCCGCTCATCGACGGCTCCGGCCTCTTCTACAAGCGCTCCAGGC 162
QY 162 CGGCAACGCTGGGAGACAGAGGTGCGCTTCTATGAGGCGTTCCTCCGCCACCGCGCGGT 221
Db 163 CGGCAACGCTGGGAGACAGAGGTGCGCTTCTATGAGGCGTTCCTCCGCCACCGCGCGGT 222
QY 222 CCGGCGCGCATCCGAGACACTTTCTTCCCGGTTCCACGGCAGCGACTCTCTCCCGAC 281
Db 223 CCGGCGCGCATCCGAGACACTTTCTTCCCGGTTCCACGGCAGCGACTCTCTCCCGAC 282
QY 282 CGAGGCGAGCGCGGGAGCGCATCTCTACCTGCTCTCGAGACCTCTCGGGGTT 341
Db 283 CGAGGCGAGCGCGGGAGCGCATCTCTCTCTCTCGAGACCTCTCTCGGGGTT 342
QY 342 TCAGGCGCCTCGCTCGCAGACATCAAGATCGGGGCCATCACTGGGCCACCGAGTTGCC 401
Db 343 TGAGGCGCCTCGCTCGCAGACATCAAGATCGGGGCCATCACTGGGCCACCGAGTTGCC 402
QY 402 GGAGCCCTACATGCGCAAGTGCCTGCGCAAGACCGCGGACCGAGCGTTCTGCTCG 461
Db 403 GGAGCCCTACATGCGCAAGTGCCTGCGCAAGACCGCGGACCGAGCGTTCTGCTCG 462
QY 462 ATTCCGCGTCTCCGCGTCCGAGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 521
Db 463 ATTCCGCGTCTCCGCGTCCGAGTCTGCTCCCGAGGCGCGCGCGCGCGCGCGCG 522
QY 522 CCGGAGGTGAAGGCCATGAGACACCGCGCGGTCTCCCGCGCGGTCTCCCGCGCTACGTGTC 581
Db 523 CCGGAGGTGAAGGCTATGAGACACCGCTCGCGCGGTCTCCCGCGGTCTCCCGCGCTACGTGTC 582
QY 582 ATCCG-TTCCGACAGAGGGATGGACTGTGCGCTCGCGCGCGGTGTACGAGGAAAG 640
Db 583 ATCCGTTTCCGACAGAGGGATGGACTGCGCGCTCGCGCGCGGTGTACGAGGAAAG 642
QY 641 GTGGAGTCTTGTACAGCTGCGGAGCTCAAGGCGTGTTCGAGGAGCAGACTCTGTTCC 700
Db 643 GTGGAGTCTTGTACAGCTGCGGAGCTCAAGGCGTGTTCGAGGAGCAGCTCTGTTCC 702
QY 701 ACTTCTACTCGCGCTCGATTTCTTCTGGCTATGATGCTGCTGCGAGTCGAGCGCGAG 760
Db 703 ACTTCTACTCGCGCTCGATTTCTTCTGGCTATGATGCTGCTGCGAGTCGAGCGCGAG 762
QY 761 ATGGGGTGGGTGACGCTGAGCTGGTGGATTTGCGCCATGTCGCCAGGAGGTGATGGG 820
Db 763 GTGGGGTGGGTAAACAGTGAAGCTGCTGCACTTTGCGCCATGTGCGCGAGGTGATGGG 822
QY 821 TGATTGACCAACTCTCTCGCGCGGTCTGCTCGCTGATCAAGTTCGTTTCTGACATTG 880
Db 823 TGATTGACCAACTCTCTCGCGCGGTCTGCTGCTGATCAAGTTCGTTTCTGACATTG 882
QY 881 TTCCGAGACTCTCTCATACGAGCCTTTGGGTCTCTTTAA 921
Db 883 TTCCAGAGACTCTCTCAGAGCGAGCCTTTGGGTCTCTTAA 923

RESULT 5
AXS13583
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
AUTHORS
Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.
TITLE
Novel inositol polyphosphate kinase genes and uses thereof
```



JOURNAL Patent: WO 02059324-A 20 01-AUG-2002;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)  
FEATURES  
source Location/Qualifiers  
1. .3416  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"  
72. .407  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52959.1"  
/db\_xref="GI:23559685"  
/db\_xref="REMBL:CAD52959"  
/translation="MPDLHPPEHOVAGHRASAKGLPLIDDSGLFYKPLQAGDRGEHE  
VAFVEAFSAHAAPARIRDTFFRFHGTLLPTEAQGEHPHLVLDLLAGFEAPCV  
ADIKIGAIT"  
ORIGIN  
Query Match 55.1%; Score 740.4; DB 6; Length 3416;  
Best Local Similarity 90.2%; Pred. No. 3.1e-102;  
Matches 862; Conservative 0; Mismatches 31; Indels 63; Gaps 4;  
QY 2 CACAGGTGAGTCTGCTGACCCCTCGCGCCATAGTCCCTTCCCGATACGATGTCGAC 61  
DB 28 CACCATGCGCCACCGCTCACCCCTTGCTCCCATAG-----TCCCGATACCATGCGCGAC 81  
QY 62 TCCACCCGCGGAGCACCAAGTCGCGCCACCGGCTCCGCGCAGCAAGCTGGGCGCG 121  
DB 82 TCCACCCGCGGAGCACCAAGTCGCGCTACCGCGCTCCGCGCAGCAAGCTGGGCGCG 141  
QY 122 TCATCGACGCGTCCGGCTCTTCTACAGCGCTCCAGCGCGGCGGACCGTGGGAGCAG 181  
DB 142 TCATCGACGACTCTGGCTCTTCTACAGCGCTCCAGCGCGGCGGACCGTGGGAGCAG 201  
QY 182 AGTTCGCGTCTTATGAGCGTTCCTCGGCGACGCGCGCTCCGCGCGCGCATCCGAGACA 241  
DB 202 AGTTCGCGTCTTATGAGCGTTCCTCGGCGACGACTCTCTCGCGGCGTTCGAGCGCG 261  
QY 242 CTTCTTCCCGCGTTCACGCGACGCGCTCTCGCGGCGTTCGCGCGCGCGCGCGCGGAGC 301  
DB 262 CTTCTTCCCGCGTTCACGCGACGCGCTCTCGGCGCGTTCGCGCGCGCGCGCGCGGAGC 321  
QY 302 CGCATCTCAGCTCTCTCGACGACTCTCTCGCGGCGTTCGCGCGCGTTCGCGCGCG 361  
DB 322 CGCATCGCACTCTCTCGACGACTCTCTCGCGGCGTTCGCGCGCGTTCGCGCGCG 381  
QY 362 ACATCAAGATCGGCGCATACGCTGCGGCGCGCGGCTTCGCGCGCGCGTTCGCGCGCAAGT 421  
DB 382 ACATCAAGATCGGCGCATACGCT----- 405  
QY 422 GCCTCGCAAGGACCGCGGCGACGAGCGTTCGCTCGGATTCGCGGCTCTCGGCGTCC 481  
DB 406 -----GACACAGAGCGTTCGCTCGGATTCGCGGCTCTCGGCGTCC 446  
QY 482 GAGTCGTGCGCGCGCGCGCGCGTGTGGCGGACGAGCGCGCGCGCGCGCGCGCGCG 541  
DB 447 GAGTCGTGCGCGCGCGCGCGCGTGTGGCGGACGAGCGCGCGCGCGCGCGCGCG 506  
QY 542 ACACCGCGGCGTCCGCGCGTCTCGGCGCTACGCTGTCATCGG-TTGGCGAGCGGG 600  
DB 507 ACATGTCGCGGCTCCGCGCGTCTCGGCGCTACGCTGTCATCGCTTCGCGAGCGGG 566  
QY 601 ATGACTGTGCGTCCGCGCGGCTGTACGAGGAGAAAGTGGAGTCTGTGTCACAGCTG 660  
DB 567 ATGACTGTGCGGCTCCGCGCGGCTGTACGAGGAGAAAGTGGAGTCTGTGTCACAGCTG 626  
QY 661 CGGAGCTCAAGGCGTGTGTGAGGAGCAGACTGTGTTCCACTTCTACTCGGCGTCCGATT 720  
DB 627 CGGAGCTCAAGGCGTGTGTGAGGAGCAGACTGTGTTCCACTTCTACTCGGCGTCCGATT 686  
QY 721 CTTCTGGGCTATGATGCTGTGTCAGTTCGACGAGCGCGAGATGGGGTGGGCTGACGGTG 780  
DB 687 CTTCTGGGCTATGATGCTGTGTCAGTTCGACGAGCGCGAGGTTGGGGTGGGCTAACAGTG 746

QY 781 AAGCTGGTGGACTTTGCCCATGTGCGCGAGGGTGATGGGGTGATTGACCACAACTTCCTG 840  
DB 747 AAGCTGGTGGACTTTGCCCATGTGCGCGAGGGTGATGGGGTGATTGACCACAACTTCCTG 806  
QY 841 GCGGGGCTCTGCTCGCTGATCAAGTTCGTTCTCACATTTGTTCCGAGACTTCCTCATACG 900  
DB 807 GCGGGGCTCTGCTAGCTGATCAAGTTCGTTCTCACATTTGTTCCGAGACTTCCTCATACG 866  
QY 901 CAGCCTTTGGTCTCTTCTTAAGAGAGATCTCGCA-TTTCGATTTGATAACAAG 955  
DB 867 CAGCCTTTGGTCTCTTCTTAAGAGAGATCTCGCA-TTTCGATTTGATAACAAG 922  
RESULT 6  
AX513578  
LOCUS AX513578 899 bp DNA linear PAT 05-OCT-2002  
DEFINITION Sequence 15 from Patent WO02059324.  
ACCESSION AX513578  
VERSION AX513578.1 GI:23559679  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1  
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.  
TITLE Novel inositol polyphosphate kinase genes and uses thereof  
JOURNAL Patent: WO 02059324-A 15 01-AUG-2002;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)  
FEATURES  
source Location/Qualifiers  
1. .899  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"  
89. .424  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52958.1"  
/db\_xref="GI:23559680"  
/db\_xref="REMBL:CAD52958"  
/translation="MPDLHPPEHOVAGHRASAKGLPLIDDSGLFYKPLQAGDRGEHE  
VAFVEAFSAHAAPARIRDTFFRFHGTLLPTEAQGEHPHLVLDLLAGFEAPCV  
ADIKIGAIT"  
ORIGIN

Query Match 51.0%; Score 686; DB 6; Length 899;  
Best Local Similarity 88.7%; Pred. No. 5.9e-94;  
Matches 802; Conservative 0; Mismatches 40; Indels 62; Gaps 3;  
QY 2 CACAGGTGAGTCTGCTGACCCCTCGCGCCATAGTCCCTTCCCGATACGATGTCGAC 61  
DB 45 CACCATGCGCCACCGCTCACCCCTTGCTCCCATAG-----TCCCGATACCATGCGCGAC 98  
QY 62 TCCACCGCGCGGAGCACCAAGTCGCGCGCACCGCGCTCCGCGCAGCAAGCTGGGCGCG 121  
DB 99 TCCACCGCGCGGAGCACCAAGTCGCGCGCTCACCGCGCTCCGCGCAGCAAGCTGGGCGCG 158  
QY 122 TCATCGACGCGTCCGGCTCTTCTACAAGCGCTCCAGCGCGCGCGCGCGCGAGCAAG 181  
DB 159 TCATCGACGCGCTCGGCTCTTCTACAAGCGCTCCAGCGCGCGCGCGCGAGCAG 218  
QY 182 AGTTCGCGTCTTATGAGCGTTCCTCGCGCAACGCGCGCTCCGCGCGCGCGCATCCGAGACA 241  
DB 219 AGTTCGCGTCTTATGAGCGTTCCTCGCGCAACGCGCGCTCCGCGCGCGCGCATCCGAGACA 278  
QY 242 CTTCTTCCCGCGTTCGAGCGCACCGGACTCTCTCCCGCAGCGGCGCGCGCGCGGAGC 301  
DB 279 CTTCTTCCCGCGTTCGAGCGCACCGGACTCTCTCCCGCAGCGGCGCGCGCGGAGC 338  
QY 302 CGCATCTCAGCTCTGCTCGACGACTCTCTCGCGGCGTTCAGGGCGCGCTGCGTCCGAG 361

```
Db 339 CGCATCGTACCTGCTCTCGACGACCTCTCCCTCGCGGGGTTTGAGGGCGCCCTGCGTCGCGAG 398
QY 362 ACATCAAGATCGGCGCCATACGTGGCCACCACCAAGTTCGCGGAGCGCTTACATCGCCAAAGT 421
Db 399 ACATCAAGATCGTSCCATCAGT----- 422
QY 422 GCCTGCCCAAGACCGCGGGACCAAGCGGTCTCTCGATTCGGCTCTCGCGCGTCCCGCGTCC 481
Db 423 -----GACCATGACGATCTGCTCGGATTCACGCTCTCCGCGTCC 463
QY 482 GAGTCTGTCGCGCCCGAGGCGCGGTGTGGCGACCGAGCGCGCGGAGGTGAAGGCATGCG 541
Db 464 GAGTCTGTCGCGCCCGAGGCGCGGTGTGGCGACCGAGCGCGCTGAGGTGAAGGCTATGG 523
QY 542 ACACCGCGCGGTCTCGCGCGCGTCTCGCGCGCTACGTGTCAFCG- TTGCCGA CGAGGGG 600
Db 524 ACATTGTGCGGCTCCGCGCGTCTCGCGCGGTGTCATGTCATCCGCTTTCGCGCGAGGGG 583
QY 601 ATGGACTGTGCGTCCGCGCGCGGTGTACGAGGAGAAAGGTGAGTCTTGTCAAGCTG 660
Db 584 ATGGACTGTGCGTCCGCGCGCGGTGTACGAGGAGAAAGGTGAGTCTTGTCAAGCTG 643
QY 661 CGCGAGCTCAAGCGCGTGTTCGAGGAGCAGACTCTGTTCACCTTCTACTCGCGCGTCTGATT 720
Db 644 CGCGAGCTCAAGCGCGTGTTCGAGGCGCAGACTCTGTTCACCTTCTACTCGCGCGTCTGATT 703
QY 721 CTTCTGGGCTATGATGCTGCTGCTGACGTGCGACGCGGAGATGGGGGTGGGGTGAAGCTG 780
Db 704 CTTCTGGGCTATGATGCTGCTGCTGCTGACGTGCGACGCGGAGGTGGGGGTGAAGCTG 763
QY 781 AAGCTGGTGGACTTGGCCATGTGCGCGAGGTGATGGGTGATGACCAACCTTCCTG 840
Db 764 AAGCTGGTGGACTTGGCCATGTGCGCGAGGTGATGGGTGATGACCAACCTTCCTG 823
QY 841 GCGCGGCTCTGCTGCTGATCAAGTTCGTTCTGACATTTGTCGCGAGACTCTCATACG 900
Db 824 GCGCGGCTCTGCTGCTGATGATCAAGTTCGTTCTGACATTTGTCGCGAGACTCTCATACG 883
QY 901 CAGC 904
Db 884 CAGC 887

RESULT 7
AP005749/c
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
BAC clone:OSUNBa0047A17, complete sequence.
ACCESSION
AP005749
VERSION
AP005749.2 GI:34740257
KEYWORDS
HTG.
SOURCE
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
REFERENCE
Sasaki,T., Matsumoto,T. and Katayose,Y.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
clone:OSUNBa0047A17
Published Only in Database (2002)
2 (bases 1 to 132170)
Sasaki,T., Matsumoto,T. and Katayose,Y.
Direct Submission
Submitted (18-SEP-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasakia@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Sep 15, 2003 this sequence version replaced gi:23200608.
The orientation of the sequence is from M13rev to -21M13 of the BAC
clone.
FEATURES
Location/Qualifiers
```

```
source
1. 132170
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="OSUNBa0047A17"

ORIGIN
Query Match 40.7%; Score 546.4; DB 8; Length 132170;
Best Local Similarity 77.5%; Pred. No. 3.1e-73;
Matches 718; Conservative 0; Mismatches 191; Indels 17; Gaps 4;

QY 13 TCCGTCACCTCTCGGCGCCATAGTCCCTTCCCATACCATAGTCCGACCTCCACCCGCG 72
Db 1983 TCCACCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1926
QY 73 GAGCACCAAGTCGCGCGCGCCACCGCGCGCTCCGCGACGAGCTGGCGCGCGCTCATGACGCG 132
Db 1925 GAGCACCAAGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1866
QY 133 TCCGCGCTCTTCTAAGACCGCTCCAGCGCGCGCACCGCTGGGGAGCAGAGGTGCGCTTC 192
Db 1865 GAGGGGCTCTTCTAAGACCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 1806
QY 193 TATGAGGGTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 252
Db 1805 TACCGCGGTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1746
QY 253 CGGTTCCACGCGCACCGGACTCTCTCCCAACGAGCGCGAGCGC---CGGGAGCGCGATCCT 309
Db 1745 CGCTTCCACGCGCACCGGCTTCTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1686
QY 310 CACTCGTCTCTGACGACCTCTCTCGGGGGTTCAGGGCGCGCTCGCTCGCAGACATCAAG 369
Db 1685 CACATCGTCTCTGACGACCTCTCTCGGGCGCTCTCCGCTCCCGCTCGCTCGCGCGAGT 1626
QY 370 ATCGCGGCCATCAGTGGCGCACCGAGTTCGCGGAGCGCGCTACATCGCGCGCGCGCG 429
Db 1625 ATCGCGCGCTGACGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1566
QY 430 AAGGACCGCGGACACGAGCGTTCGTCTCGGATTCGCGCGTCTCGCGCGTCCGAGTCTGC 489
Db 1565 AAGGACCGCGGACACACGCGCGCTCTCTCGGCTTCGCGCTCTCGGCGTCTCGGGTGTG 1506
QY 490 ---GGCGCGGAGCGCGCGTGTGGGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
Db 1505 GATGCCCGGGCGCGCGCGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1446
QY 547 GCGCGCGTCCGCGCGCGTCTCCGCGCTACGTGTATCGTTCGCGCGAGCGGCGATGAC 606
Db 1445 GCGCGCGTCCGCGCGCGTCTCCGCGCTACGTGTTCGCGCGCGCGCGCGCGCGCGCGCG 1386
QY 607 TGTGCGCTCGCGCGCGCGGTGTACGAGGAGAAAGGTGGAGTCTTGTCAAGCTCGCGGAG 666
Db 1385 TGTGCGCTCGCGCGCGCGGTGTACGAGGAGGAGGCGCGCGCTTGGCTCAGCTGCGGAG 1326
QY 667 CTCAGGCGTGTGTGAGGAGCAGACTCTGTTCACATTTCTACTCGCGCGTCTGATTTCTTG 726
Db 1325 CTCAGGCGTGTGTGAGGAGCAAAACCTGTACCACTTCTACTCTCGCGCGTCTGATTTCTG 1266
QY 727 GGCT-----ATGATGCTGTGCGTTCGAGTCGACGCGCGCGGAGATGGGGGTGGGTGACG 777
Db 1265 GGCTACGACGCGCAATGCGCGCGCGCGCGCGCTGCTCCGAGAGTGGAGCGCGGTGAAGG 1206
QY 778 GTGAGCTGTGGAGCTTTTTCGCCATGTGGCGCGGAGGTGATGGGTGATGACACAACTTC 837
Db 1205 GTGAGCTGTGGAGCTTCGCGCATGTGACGATGCGGAGCGGGGTGATGACACAACTTC 1146
QY 838 CTGGCGCGGCTCTGCTGCTGATCAAGTTCGTTCTGACATTTCTCGGAGACTCTCTCAT 897
Db 1145 TTGGCGCGGCTCTGCTGCTCATCAAGTTCATCGCGCGACATTTGTGCGCGGATTACCAG 1086
```



```
QY 547 GCGGCGTCCGCGCGTGTCTCGGCGCTACGTGTCTATCGTTGCCGACGAGGGGATGGAC 606
Db 882 GCGGGGTCCGCGCGTGTCTCGCGCTACGTGTCCACGGCGCGGCGACGGCCTGGAC 941
QY 607 TGTGCGTCCGCGCGCGTGTACGAGGAAAGGTGGAGTCTTGTCAAGCTCGCGAG 666
Db 942 TCGCGCTCCGCGCGCGTGTACGAGGGAGGGCGCGTCTGTGCTAGCTCGCGGAG 1001
QY 667 CTCAGGCGTGTTCGAGGAGCAGACTCTGTTCACCTTCTACTCGCGCTCGATTCTTCG 726
Db 1002 CTCAGGCGTGTTCGAGGAGCAAACTGTACCTTCTACTCGCGCTCGATTCTGTTC 1061
QY 727 GCGT-----ATGATGCTCTGAGTCGAGCAGCGCGGAGATGGGGTGGGTGAGG 777
Db 1062 GCGTACGACGCAATGCGCGCGCGCGTCTCGCGAGGTGGAAGCGCGGTGTAAGG 1121
QY 778 GTGAAGCTGGTGGACTTGTGCCATGTGGCGCGAGGTGATGGGTGATTGACCACAACTTC 837
Db 1122 GTGAAGCTGGTGGACTTGTGGCGATGTGACGATGGGACGGGTGATTGACCACAACTTC 1181
QY 838 CTGGCGCGGCTGTGTGCTGTGATCAAGTTCGTTTCTGACATGTTCCGAGACTCCTCAT 897
Db 1182 TTGGCGCGGCTGTGTGCTGTGATCAAGTTCATCGCGACATGTCGCGAGGTACCAG 1241
QY 898 AGCAGCCTTGTGGTCTCTTTAAGA 923
Db 1242 AAGCGCTCTCAGATCATCTTTGAAA 1267

RESULT 9
AX513580
LOCUS AX513580 643 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 17 from Patent WO02059324.
ACCESSION AX513580
VERSION AX513580.1 GI:23559681
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 17 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source
1. .643
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 35.0%; Score 470.4; DB 6; Length 643;
Best Local Similarity 95.9%; Pred. No. 1.9e-61;
Matches 509; Conservative 0; Mismatches 18; Indels 4; Gaps 3;
QY 42 TCCCATACCATGTCGACCTCCACCGCGGAGCACCAGTCGCGGCCACCGCGCTC 101
Db 107 TCCCATACCATGTCGACCTCCACCGCGGAGCACCAGTCGCGGCCCTC 166
QY 102 CGCAGCAAGCTGGGCGCGCTCATGACGGCTCCGGCTCTTCTACAGCCGCTCAGGC 161
Db 167 CGCAGCAAGCTGGGCGCGCTCATGACGGCTCCGGCTCTTCTACAGCCGCTCAGGC 226
QY 162 CGGCAACCGTGGGAGCAGAGTGCCTTCTATGAGGGTTCCTCGCCACCGCGCGT 221
Db 227 CGGCAACCGTGGGAGCAGAGTGCCTTCTATGAGGGTTCCTCGCCACCGCGCGT 286
QY 222 CCGGCGCGCATCCGAGACACCTTCTTCCCGGTTTCCACGGCAGCGACTCCTCCCGAC 281
Db 287 CCGGCGCGCATCCGAGACACCTTCTTCCCGGTTTCCACGGCAGCGACTCCTCCCGAC 346
```

```
QY 282 CGAGCGCAGCCCGCGGAGCCGCACTCTCACTCTCGACGACCTCTCTCGCGGGTT 341
Db 347 CGAGCGCAGCCCGCGGAGCCGCACTCTCGCACTCTCTCGACGACCTCTCTCGCGGGTT 406
QY 342 TCAGCGCCCTGCTCGCGAGACATCAAGATCGGCGCATCACTGTCGCCACCGAGTTCCGC 401
Db 407 TGAGCGCCCTGCTCGCGAGACATCAAGATCGGCGCATCACTGTCGCCACCGAGTTCCGC 466
QY 402 CGAGCCCTTACATCGCCCAAGTCTCTCGCCAGGACCGCGGACCAAGAGCTTCTCTCGG 461
Db 467 GGAGCCCTTACATCGNCAAGTACTTNGCCAGGACCGCGGACCAAGAGCTTCTCTCGG 526
QY 462 ATTTCGCGCTCTCGCGCTCTCGAGTCTCGGCGCGGCGCGCTGTCGCGGAGCGAGG 521
Db 527 ATTTCGCGCTCTCTCGCGCTCTCGAGTCTCGGCGCGGCGCGCTGTCGCGGAGCGAGG 584
QY 522 -CCCGAGGTCAAGCCATGGACA-CGCGCGCGCTCTCGCGCGCTCTCGG 570
Db 595 CCCCGGGGTGAANGCTATGGACACCCGTCGNGNCCGCGGCTCTCTCGG 635

RESULT 10
AX513581
LOCUS AX513581 519 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 18 from Patent WO02059324.
ACCESSION AX513581
VERSION AX513581.1 GI:23559682
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 18 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source
1. .519
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 22.2%; Score 298.4; DB 6; Length 519;
Best Local Similarity 89.4%; Pred. No. 1.6e-35;
Matches 328; Conservative 0; Mismatches 38; Indels 1; Gaps 1;
QY 627 GTACGGAGAAAAGGTGGAGTCTTGTACAGCTCGCGAGCTCAAGCGTGGTTGAGGA 686
Db 2 GTACGGAGAAAAGGTGGAGTCTTGTACAGCTCGCGAGCTCAANGCGTGGTTGAGGG 61
QY 687 GCAGACTCTGTTCACCTTCTACTCGCGTCTGATTTCTCTGGGCTATGATGCTGCACT 746
Db 62 GCAGACTCTGTTCACCTTCTACTCGCGTCTGATTTCTCTGGGCTATGATGCTGCACT 121
QY 747 CGCAGCAGCGGAGATGGGGTGGGTGAAGCTGGTGAAGCTGGTGAAGCTTTCCTATGTCG 806
Db 122 CGCAGCAGCGGAGATGGGGTGGGTGAAGCTGAAGCTGGTGAAGCTTTCCTATGTCG 181
QY 807 CGAGGTGATGGGTGATGATGACCAACTTCTCGGGGGGCTCTGCTCGCTGATCAAGTT 866
Db 182 CGAGGTGATGGGTGATGATGACCAACTTCTCGGGGGGCTCTGCTCGCTGATCAAGTT 241
QY 867 CGTTTCTGACATGTTCTCGGAGACTCTCTACACAGCCCTTTGGTCTCTTTAAGAGAG 926
Db 242 TGTTCGACATGTTCTCAGAGACTCTCTCAGACCGAGCTTTGGTCTCTTTAAGAAAA 301
QY 927 GATCCTGSCA-TTTCGATTTGATAACAAAGCCCTTACAGTTTTCCTTGGAAAAAGAGCG 985
Db 302 GATCCTGSCATTTTCGATTTGATAACAAAGCCCTTACAGTTTTCCTTGGAAAAAGAGCG 361
```

```
QY 986 CCTCCGA 992
Db 362 CCAGTGA 368

RESULT 11
AX513582
LOCUS AX513582 353 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 19 from Patent WO02059324.
ACCESSION AX513582
VERSION AX513582.1 GI:23559683
KEYWORDS
SOURCE
ORGANISM
Eucalyptus grandis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 19 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
1..353
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN
Query Match 14.3%; Score 191.6; DB 6; Length 353;
Best Local Similarity 90.2%; Pred. No. 2.1e-19;
Matches 238; Conservative 0; Mismatches 21; Indels 5; Gaps 3;
QY 667 CTCAGGCGTGGTTCGAGGAGCAGACTGTGTTCCACTTCTACTCGGCGTCGATTTCTTG 726
Db 1 CTCAGGCGATGTTGGAGGAGCAGACTGTGTTCCACTTCTACTCGGCGTCGATTTCTTG 60
QY 727 GCCTATGATGCTGCTGAGTGCAGCAGCGGAGATGGGGTGGGGTGAAGCTG 786
Db 61 GCCTATGATGCTGCTGAGTGCAGCAGCGGAGTGGGGTGAAGCTG 120
QY 787 GTGACATTTGCCATGTCGCGAGGATGATGGGT--GATTGACACACACTTCTTGGCG 844
Db 121 GTGACATTTGCCATGTCGCGAGGATGATGGGTGATTGACACACACTTCTTGGCG 180
QY 845 GCCTCTGCTGCTGATCAAGTTC--GTTTCTGACATTTCTCGGAGACTCTCTATACG-C 901
Db 181 AGCTCTGCTAGTCAAGTTCGTTTCTTGACATTTCTCGAGNACTCTCTAGACGCC 240
QY 902 AGCCTTTGGGTCCTTTTAAAGA 925
Db 241 AGCCTTTGGGTCCTTCTTAAAA 264

RESULT 12
AX513574
LOCUS AX513574 1195 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 11 from Patent WO02059324.
ACCESSION AX513574
VERSION AX513574.1 GI:23559674
KEYWORDS
SOURCE
ORGANISM
Eucalyptus grandis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
rosids; Myrtales; Myrtaceae; Eucalyptus.
1
REFERENCE
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 11 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
1..1195
/organism="Eucalyptus grandis"
/mol_type="unassigned DNA"
/db_xref="taxon:71139"
116--1048
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52956.1"
/db_xref="GI:23559675"
/db_xref="RENTREMBL:CAD52956"
/translation="MLKVPDQVAGHRGDKGLPLVDDSGRFYKPLQSDHRGDTVEA
FYESFYNTPEPHIRKEFFPAFHCTKTIEASDGGPQPHLVLELIVSRPNSLMDIK
TGSTWPEASEEYIQKLEKDRNSTSVSLGRVYONLEAGFWQPEKXVYSF
NADGVSALIRKFVSSNLGPNVDPDCLLYASKYVCHRGGLLAQLQLKEWFEVQNYH
FYCSLLIYDRESALDGCAPKVLVDFAHVMDHGVIDHNLGGLGSLVKIFIRIDIA
DEDNKCAKCEVNLKENGFKYSTPELDEAC"

ORIGIN
Query Match 13.7%; Score 184.6; DB 6; Length 1195;
Best Local Similarity 55.6%; Pred. No. 2e-18;
Matches 471; Conservative 0; Mismatches 349; Indels 27; Gaps 5;
QY 56 CCGACCTCCACCCCGGAGCACCAAGTCGCGGCCACCGCGCTCCGCCAGCAAGCTGG 115
Db 114 CCATGCTCAAGTCCCGATCTCAAGTCGCGGTACCGGGGAGACGGGGAAAGCTGG 173
QY 116 GCCCGCTCATCGACGGCTCCGGGCTCTTTCTAACAAGCGCTCCAGCCGCGACCGTGGG 175
Db 174 GGCCACTGGTGGATGATTGGGGCGCTTCTATAAGCCTCTCCAGAGCGATCATCGCGAG 233
QY 176 AGCAGGAGTGCCTTCTATGAGCGGTCTCCGCCACCGCGCTCCGCCCGCGATCC 235
Db 234 ACACGGAAGTGGCGCTTTTACGAGTCACTTATTCNAATACCGAGATCCAGGTCACTTC 293
QY 236 GAGACACCTTCTCCCGCGGTTCACGCGACGCGACTCTCCCCACCGAGCGGAGCCCG 295
Db 294 G---CAATTTCTTCCTCGGTTTACGGAATAAGACTAT---TGAGCGCTCTGATGAT 347
QY 296 GGGAGCGCATCTTCACCTCGTCTCGACGACCTCTCGGGGGTTTCAGCGCGCTTCGCG 355
Db 348 CGGCTCTCAACCTCACCTGTTCTGAGGATCTCGTCTCGGGTCGACGAACCACTC 407
QY 356 TCGCAGACATCAAGATCGGCGCATACGTGGCCACCGAGTTCGCGGAGCCCTACATCG 415
Db 408 TCATGGGACATCAAGACTGGATCCAGAAATGATATCGGAGGCTCTGAGGAGTACATPC 467
QY 416 CCAAGTGCCTGCCAAGGACCGCGGACCGAGCGCTCTCGTCCGATTCGCGCTCTCCG 475
Db 468 AAAAGTGTAGAGAAAGATCGAAATAGCACACGCTTCAATGGGTTTATAGGATTTCTG 527
QY 476 GCCTCCGAGTGTGCG---CCCCGAGGGCGCGTGTGCGGACGAGCGCGCGGAGGTGA 532
Db 528 GGCTAAGGGTATATCAAAATAGCGAAGCTGGATTTTGGCAACCTGAGAAGAGGTTGTT 587
QY 533 AGGCCATGGACACCGCGGGTTCGCGCGGTGCTCCGGCGCTACGTGTATCC----- 585
Db 588 ATAGCTTTAATTCGCGACCGGTGTACAGTCCGGTCTGAGGAAGTTTGTTCCTTCAACTGT 647
QY 586 -----GTTGCGGACGAGGGATGACATGTGCGCTCGCGCGCGGTGTACGAGGAAAG 640
Db 648 CTCGGGGTCCAAATGTGATCCGGATTTGTTGTATGATCAATAAAGTTTACTGTACCGGG 707
QY 641 GTGAGGTCTTTGTACAGCTCGCGGAGCTCAAGCGGTGTTTCGAGGAGCAGACTCTGTTC 700
Db 708 GTGGAATTTGGCAACAATGCTTCAGCTGAAGAAATGTTTGAAGTTTTCAGACGAATATC 767
QY 701 ACTTCTACTCGGCTCGATTTCTTCTGGCTATGATGCTGTGCTGAGTCCGACGAGCGGAG 760
Db 768 ACTTCTATTCTTGTTCATCTATTATATGACAGGAGTCTGCTTTGACGCGC----- 823
QY 761 ATGGGGGTGGGTGACCGTGAAGCTGTGCTGATGCTGTGCGCGAGGTTGATGGG 820
Db 824 --TGTCGACACCCGAAAGTTAACTGGTGGACTTTGACATGTGATGATGATGATGATGATG 881
```

```
QY      821  TGATTGCCAACCTTCCTGGCGGCTCTGCTGCTGATCAAGTTCGTTTCTGACATTG 880
Db      892  TGATGATCAACATTCCTGGTGGCTCTGTTCTGTAATCAAGTTTATACGTGACATTG 941
QY      881  TTCGGA 887
Db      942  CTGATGA 948

RESULT 13
LUS310150      25054 bp      DNA      linear      PLN 24-JAN-2002
Locus          Linum usitatissimum variety Bombay Ngc-D, Ngc-A and Ngc-B genes.
DEFINITION     AJ310150
ACCESSION      AJ310150
VERSION        AJ310150.1 GI:13509206
KEYWORDS       Ngc-A gene; Ngc-A protein; Ngc-B gene; Ngc-B protein; Ngc-D gene;
               Ngc-D protein.
SOURCE         Linum usitatissimum (flax)
ORGANISM       Linum usitatissimum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Linaceae; Linum.
1
DODDS,P.N., Lawrence,G.J. and Ellis,J.G.
Contrasting modes of evolution acting on the complex N locus for
rust resistance in flax
Plant J. 27 (5), 439-453 (2001)
21461288
PUBMED 11576428
REFERENCE      2 (bases 1 to 25054)
AUTHORS        DODDS,P.N.
TITLE          Direct Submission
SUBMITTED      (22-MAR-2001) DODDS P.N., Plant Industry, CSIRO, GPO Box
1600, Canberra, ACT 2601, Australia
LOCATION/Qualifiers
1..25054
/organism="Linum usitatissimum"
/mol_type="genomic DNA"
/variety="Bombay"
/db_xref="taxon:4006"
/haplotype="N"
Join(2881..3431,3533..4652,4770..5045,5140..6519)
/gene="Ngc-D"
Join(2881..3431,3533..4652,4770..5045,5140..6519)
/gene="Ngc-D"
/note="N rust resistance gene: TIR-NBS-LRR protein"
/codon_start=1
/product="Ngc-D protein"
/protein_id="CAC35321.1"
/db_xref="GI:13509207"
/db_xref="GOA:O9ARC4"
/db_xref="SPTREMBL:O9ARC4"
/translation="MMRSDSGTSTFHSKSSADLTFFPLPSGEVEVFLSPRGDVRK
TFADHLYTSLVRSKIRTFDEEGQKGETIGPSLIQAITESKIYIPILPTQYASSKWC
LQELAKVGVKSGGARGQHIIIPVFMFDPDRVDPDPSPGPKYKESFOHNLKADPET
LLEWKALQEVGMKGWHISLTGQGAVDKIFTEVELHLRANTYLTATDELVLGDFSV
DEWKLNLDSSTSEKIIGYMGRLGKTLTATVYNKVSQFECFIDNIRETLKN
DGVALQNKVSDILRKDFCOAKNADGVOMIRERSHKKIFVVLDDVNESFRDDIF
GKLTAFSADSRFLVTTDRATLELRGCKLFKHGMSHSLKLFKSHAFQVDYPED
YASICEFVQVSGSLPALKVIGSLLPRTSKSPKDKLIELKAI PANVOYRAKISYN
ELTDNEKQIELDVACLVEFVGAKEIPIITWSDCGFYPTTITRLVQSLVIRINDEEFW
MHDRIQGLAIQVLEVLNGDLGNFNVLPSLEFRLVHYHGDGCPGSLNKLMLLEL
LTVNEFQTNIRLPSLEVLNGDLGNFNVLPSLEFRLVHYHGDGCPGSLNKLMLLEL
EVSVDTWEGNREIKAAKGLKVHLMCKLEKVPDLSTCRLELLRFSICRMHGE
LDINFKDLVLTDFQIRITALKEVESLQNLQDLVSSGLIEVPAGISKLSSLYL
LNTNKHDKVETLPNGLKILLISFSLSPSSIFRLDVRYSWLRPLNLSVTNLT
RLRUEVGHIGPGLSELKLECLFLDAENLNDGLENLVLKELAVECRILEKL
PSTALETKHKLITGQCNILGIBYGLANLGSLSHLEISGCPCLTVVESLHSLNLGT
LELSGYGTNITLPPSLITYTKLSKVSQSLQDLPTNLKNLRCIKICGDNFIFITGL
HTSLEELRVMSGSIRKDLTLGLVLEILQFDSCTQITRIGLGLESQRLHMSRC
QSIKELPNLSGLKILSYILLEKCRHLKEVYRLBELKRLDFNTHRG"
17600..18457
/function="unknown"
/note="ORF
related to human inositol hexakisphosphate kinase 2
protein"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAC35324.1"
/db_xref="GI:13509210"

variation
```



```
/db_xref="GOA:Q9ARCl"
/db_xref="SPTREMBL:Q9ARCl"
/translation="MLKVPHEOVAGHGDINGLLPLVDDSGRFYKPLQDDEGATEMA
FTSTSTNTKIPQXVRRPFOAQFGQSLSEASDGLHPLHLLDITSLHLCVMDIK
IGSRWYEPASQAYIEKCLKUDVESNELLPFRISGLQYVNGKSESSEVILKPERKLL
LNLNADEVLLKRVFNSNLSKSDPQCSFAAVVSSGSLIAQLLELKAWEFQDTIYH
FISNQIML"
join(20265..20815,21104..22202,22415..22690,22818..24218)
/gene="Ngc-B"
CDS
join(20265..20815,21104..22202,22415..22690,22818..24218)
/gene="Ngc-B"
/note="N rust resistance gene homolog: TIR-NBS-LRR
protein"
/codon_start=1
/product="Ngc-B protein"
/protein_id="CAC3325.1"
/db_xref="GI:13509211"
/db_xref="GOA:Q9AR40"
/db_xref="SPTREMBL:Q9AR40"
/translation="MWRSDTSSIGSFLSRGSDVPTLPLPSGEYEVFLSPRGPDVRQ
TFADHLVAVLRSKIRTRDEGLQKGTIGSSLIQALETSEKIVIPILTONYASKKWC
LOELAKMVDCKWNGGGGQHLILPVFPMRDPVRHDSGPYKEAFQHNKHPET
LLEWEALQDVGMKGWHEINELTGQAVVDKIFTTIEFHLRANYLATDELVGIDSSV
EVEMLMDHDSERIIGIYGMGLGKTTAKAVFNKVSQFVCCDDINIRETLR
NDGVVALQNVISDILRDSQAKNASDGVRIIRVRHKIFVYLDIDIESFHFDEI
FGKLGDSFDSRFLITRDARTLELLNECKMFGLEMSHDSLOLFSKHAQVDPPE
DYASICEEFIVASGLPLALVIGSLFKDKRPWEDKILKAIKPSAKVQERLKVSY
NELTHNEKOIFLDIACLVAKKEVPMYMSDCLYPASTRTLVORSLVRMDNKKF
WMHDIRLQRTIVRENSQNPYKRSRWSNNDALDIILKNREGNDVVALRDMKGG
YALINKEFKSRFLNFLENGDLKSNFNILPNLRWLVYRGDPSPGLNKLNLVIL
ELDGYVTHSWKWNIEKAAGLKVNTSGILEKVPDLSTCRLELLCFHKCOMMR
GELDTGFKDLKVDINOTETTLKGVNESLQNLQDLVGRSLIEVPAGISKLSSLE
FLDITSVKHDEVEMLPNGILVPSFSLPSLILKDI CDSRNLQRLNLASVTVN
LTRHLKVGHEIPELGLKLLESLSICNAPNLNDLGNLVLKELALRCPILG
KLPSAEITKLHKVIRCDVLGEIYGLGNLSDLSHLDISWCPRLTVMDDLHLLKL
GTLVSSGELTNLPSLUSITKRLTEVRSQPDNLNKNRDLITITGRELIEFIA
GLHTLESUELSEMERCPVRKIDLAGLKJTHICTQLTETRLGLGSLQMLFIA
SGCOSIKELPNLGLNKLKFLSKERQKLVNGLLELEWLDFTDRRLKLYLLKLV
SRKGQLVTVQARRALFINN"
ORIGIN

Query Match 13.5%; Score 181.6; DB 8; Length 25054;
Best Local Similarity 55.4%; Pred. No. 3.6e-18;
Matches 474; Conservative 0; Mismatches 349; Indels 33; Gaps 5;

QY 50 CCAUTGCGACCTCCACCGCGGAGACCAAGTCGCGGCGCACCGCCTCGGCCACA 109
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 9605 CCAAGCCATGTTCAAAGTTCGGAACTCAGTTGCTGGACATCAAGCCATCAACGCC 9664
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 110 AGCTGGGCGCGTCATCAGCGCTCGCGCTCTTCTACAAGCGCTCCAGCGCGCGGACC 169
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 9665 TCCTCGGTCTTCTAGACGATTCAGGCGGATTCACAGCCATTCAGGCGGATGAC 9724
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 170 GTGGGAGACAGAGTCGCTCTATAGAGCGTTCTCCGCCACCGCCCGCTCCCGGCC 229
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 9725 GTGGGCAACCGAGATGCAATCTACATCTTTTCTCTGCTACAAAACGATTCACAAC 9784
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 230 GCATCCGAGACACCTCTTCCCGGTTCACGGCACCGGACTCTCCCCACCGAGCGC 289
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 9785 AGTCCG---GAGGTTTTCGGGCTTCATGGGACTCAAGTATAGAGGCTTCGTATG 9841
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 290 AGCCGGGAGCGCATCTCATCTCGTCTCGACGACTCTCTCGCGGGTTTCAGGCGC 349
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 9842 GATCTGG---CCTCACCGGACCTCATCTTAGAAGATCTCAGCGTCACTCGGTGATC 9898
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 350 CTGCGTCGACATCAAGATCGCGGCATCAGTGCGCACCGAGTTCCGCGGAGCCCT 409
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 9899 CATGTGTATGACATCAAGATTTGGTTCCAGGACATGTGTATCCCGAGGCTTCTCAAGCCT 9958
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 410 ACATCGCAAGTGCCTCGCCAGGACCGCGGACACAGCGTCTCTGCTCGGATCCGCG 469
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 9959 ACATCGAAGATGCTGAGAGGATGTGGAATCAAGCAGTCCCTTCTCGGTTTAGGA 10018
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 470 TCTCCGCGTCGAGTCTGTCGSCCCCGA-----GGCGCGCTGTGCGCGACGG 517
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10019 TATCCGGATTGAGGTTTACGTAACGACAAAGAATCTCTGAGGTGGTCTTTGAAGCCTG 10078
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 518 AGCGCCGCGAGGTGAAGCCCATGACACCGCGCGGTCCGCGCGTGTCTCCGCGCGTACG 577
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10079 AGAGGAAGCTTCTCCAGAACTCTTACTGCTGATGAAGTTAGGCTGGTCTTAAACGTTTG 10138
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 578 TGTCAATCCGTTGCCGA---CGAGGGGATGAGACTGTGCGCTCCGCCGCGGTTGACGAG 634
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10139 TTTCTCCAAACCGGAAGTCGACCAACCGGATTTGTTGCTTTCAGCGGTGCTTTACGGTG 10198
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 635 GAAAGAGTGGAGTCTTGTACAGCTGCGGAGCTCAAGCGGTGTTCCGAGGAGCAGATC 694
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10199 GTTCAATGGGATCTTGGCTCAGTTCTGAGCTGAAGCGGTGGTTGAGGATCAGACA 10258
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 695 TGTTCACCTTCTACTCGGCTCGATTTCTTCTGGGCTATGATGCTGCT-----G 742
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10259 TCTATCATTTCAACTCTTGTCTTCTTATGCTATGAGAAGGAGAAGAGATGG 10318
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 743 CAGTCGAGCAGCGGAGATGGGGTGGGTGACAGCTGAGCTGTGACTTGTCCCATG 802
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10319 TGAATGGAGGAGAGAGAGTTTGGGTGCTGTCGGGTGAAGCTCATCGATTTTGTCTCATG 10378
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 803 TGGCGAGGCTGATGGGTGATTGACCAACTTCTCGCGCGGCTCTGCTCGCTGATCA 862
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10379 TGACGAGGAAATGGCGTTATCGATCACAATTTCTTGGAGGGCTCTGCTCATTGATAA 10438
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 863 AGTTCGTTCTGACAT 878
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10439 AGTTCATTTCCGAGAT 10454
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
AX513572 1105 bp DNA linear PAT 05-OCT-2002
LOCUS AX513572
DEFINITION Sequence 9 from Patent WO02059324.
ACCESSION AX513572
VERSION AX513572.1 GI:23559672
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1
AUTHORS Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 9 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source
1..1105
/organism="Glycine max"
/mol_type="unassigned DNA"
/db_xref="taxon:3847"
12..851
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52955.1"
/db_xref="GI:23559673"
/db_xref="REMBL:CAD52955"
/translation="MLKIPHEOVAGHAKDGLPLVDDFGKFKYKPLQNKDDTRGS
TELSFYSLAAAHDSIRSFPFAPHGTTRLLDSDGSPHPHVLIEDLLCGYKPSVM
DWKIGATWHLGSDSEYICKLKADRESSLPIGRISGVKXDSISSMEPTKRSIQCLS
AHKGLVNLNKFVSNINDDHDPDFACAFTEVYAGVLERLQKLDKWFQVQVHYFSC
SLVWVYKDLGKATNPLVKLVDFAHVVDGNGVIDHNFLLGSLCFKFKLKDILAVC
LHK"
ORIGIN

Query Match 13.3%; Score 179; DB 6; Length 1105;
Best Local Similarity 55.6%; Pred. No. 1.4e-17;
```



Matches 460; Conservative 0; Mismatches 340; Indels 27; Gaps 5;

QY 61 CTCACCCCGCCGAGACCAAGTCGCGCCGACCGCTCGCCAGCAAGCTGGGCGG 120

Db 15 CTCAGATCCCGAGACCAAGTGGCGCGGCAAGGCCAAGACGAATCTGGGCCA 74

QY 121 CTCATCGAGCGTTCGGGCTCTCTTACAAAGCGCTCCAGCGCGGACCGTGGGAGCAC 180

Db 75 CTCGTGACGATTTTGGAAATTTACAAAGCCCTCCAGACCAAAAGACGACAC 134

QY 181 GAGTGCCTTCTATGAGCGTT-----CTCGGCCACGCGCGCTCCGCCCGCATC 234

Db 135 CGCGGCTCCACCGAATCTCCTTTACACCTCTCTCGCGCGCGCCACGACTACTC 194

QY 235 CGAGACACCTTCTTCCCGCGTTTCCACGCGACGCTCTCTCCACCGAGCGCGCC 294

Db 195 ATCCGCTCTCTTCCCGCGTTTCCAGGCAACCGCTCTCGACGCTCCAGCGCTC 254

QY 295 GGGAGCGCGATCTCTACCTCGCTCGACGCTCTCTCGCGGGTTTCAGCGCCCTGC 354

Db 255 GG---TCCCGACCTTCACTGTCTCTGGAGGACCTCTCTCGCGGTACTCCAAAGCCCTCC 311

QY 355 CTCGACGATCAAGATCGCGCCATCAGTGGCCACCGAGTTCGCGGAGCCCTACATC 414

Db 312 GTCATGGAGTAAGATCGGCTTCCAGACCTGCGACCTGGAGACTTCGAGGACTACATC 371

QY 415 GCGAAGTCTCGCCAAAGACCGCGGACCGAGGTTCTGCTCGATTCGCGGCTCTCC 474

Db 372 TCGAAGTCTCGAAGAGACAGAGTCTCTAGCTTGCCTTGGTTTCAGAACTCTCG 431

QY 475 GCGTCCGAGTCTCGCGCCCGAGGCGCGCTGTGGCGGACGAGCGCCCGAGGTGAAG 534

Db 432 GGAGTCAAGGACTCTATCTCTCTCTGGAACTTACAGGAAATCTCTCAGTGTCTATCC 491

QY 535 GCGATGGACACCGCGGCTCGCGCGCTGCTCGCGGCT---ACGTGTCATCCGTTGCC 591

Db 492 GCGCATGGTGTGCACTTGTCTCAACAGTTCGTTTCTCTAATAATATCAACCATGAT 551

QY 592 GACGAGGAGTGAAGTGTGCTCGCGCGCGGCTGTACGGAGGAAAGGTGGAGTCTTG 651

Db 552 GATCATATCCGATTCGCTTTCGCAACGGAGTCTAC-----GGCGCGCTTTG 602

QY 652 TCACAGTCCGCGAGCTCAAGCGTGTTCGAGGAGCAGACTCTGTTCACATTCATCTCG 711

Db 603 GAGCGCTTCGAGAAAGCTCAAGGACTGGTTCGAGGTTTCAGACGCTGTATCACTTCTATTCT 662

QY 712 CGGTGATTCCTTGGGCTATGCTGCTGAGTTCGAGTTCGAGGAGGAGATGGGCTGGG 771

Db 663 TGTTCGTCTTGTGTGTACGAGAGGATCTAGGAAA-----GGGAACTACCAAC 716

QY 772 GTGACGGTGAAGCTGGTGAAGTTGCCCCATGTGGCGAGGTTGATGGGTGATTGACCAAC 831

Db 717 CCTCTGCTCAAACTCGTTGACTTTCACACGCTGTTGACGAGGAAACGTTGCTATTGATCAC 776

QY 832 AACTTCTTGGGCGGCTCTGCTCGCTGATCAAGTTCGTTCTGACAT 878

Db 777 AACTTCTTGGGCGGCTTGTTCCTTCATCAAGTTCCTCAAGGATAT 823

## RESULT 15

AP004772  
LOCUS Oryza sativa (japonica cultivar-group) 116893 bp DNA linear PLN 03-OCT-2003  
DEFINITION PAC clone:P0415B12, complete sequence.  
ACCESSION AP004772  
VERSION AP004772.2 GI:37497094  
KEYWORDS HTG.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.

## REFERENCE

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC clone:P0415B12  
JOURNAL Published Only in Database (2002)  
REFERENCE 2 (bases 1 to 116893)  
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
TITLE Direct Submission  
JOURNAL Submitted (20-FEB-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail:tsasakienias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-7441, Fax:81-298-38-7468)  
COMMENT On Oct 2, 2003 this sequence version replaced gi:18844992.  
The orientation of the sequence is from SP6 to T7 of the PAC clone.

FEATURES  
Location/Qualifiers  
1..116893  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="2"  
/clone="P0415B12"

ORIGIN  
Query Match 11.6%; Score 156.2; DB 8; Length 116893;  
Best Local Similarity 74.7%; Pred. NO. 1.9e-14;  
Matches 213; Conservative 0; Mismatches 63; Indels 9; Gaps 1;

QY 648 CTTGTCACAGCTGCGGAGCTCAAGCGTGGTTTCGAGGAGCAGACTCTGTTCCACTTCTA 707

Db 1 CCTGGCTCAGTTCGGGAGCTCAAGCGTGGTTTCGAGGAGCAACCTGTACCACTTCTA 60

QY 708 CTCGCGCTCGATTTCTTCTGGGCTATGA-----TGCTGCTGAGTCGACGAGGCGG 758

Db 61 CTCGCGCTCGATTTCTGTTTCGCTACGACCCAAATCGCGCGGCGGCTGCTCCGAGG 120

QY 759 AGATGGGGTGGGTTGACGCTGAAGCTGCTGACCTTTGCCCATGTGGCCGAGGGTGAATGG 818

Db 121 TGAAGCGCGGTGTAAAGGTGAAGCTGCTGTCGATTCGCGCATGTCGACGATGGGACGG 180

QY 819 GGTGATTCACCACTTCTTCTGGGCGGCTGCTGCTGATCAAGTTCGTTCTGACAT 878

Db 181 GGTGATTCACCACTTCTTCTGGGCGGCTGCTGCTGATCAAGTTCATCGGCGACAT 240

QY 879 TGTTCGCGAGACTCTCTATACGCGACCTTTGGGCTCTTCTTAAGA 923

Db 241 TGTTCGCGAGGTACCGAGAGGCGCTTTCAGATCATTTCTTGAA 285

Search completed: March 27, 2004, 03:33:19  
Job time : 5360 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2004, 23:13:14 ; Search time 601 Seconds  
(without alignments)  
9500.130 Million cell updates/sec

Title: US-10-042-894A-7  
Perfect score: 1344  
Sequence: 1 gcacgagtcagtcggtcac.....ataaaaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1980s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1344	100.0	1344	6	AAD43514
2	864.4	64.3	922	6	AAD43513
3	844.8	62.9	1169	6	AAD43511
4	837.4	62.3	923	6	AAD43512
5	740.4	55.1	3416	6	AAD43522
6	686	51.0	899	6	AAD43518
7	470.4	35.0	643	6	AAD43519
8	238.4	22.2	519	6	AAD43520
9	191.6	14.3	353	6	AAD43521
10	184.6	13.7	1195	6	AAD43516
11	179	13.3	1105	6	AAD43515
12	149.8	11.1	1020	6	AAD43517
13	148.6	11.1	1168	3	AAC33023
14	147	10.9	1104	3	AAC33685
15	147	10.9	1243	3	AAC33692
16	145.6	10.8	1130	3	AAC48750
17	142	10.6	1130	3	AAC40313
18	104.6	7.8	464	6	ABL93575
19	72.4	5.4	113193	7	AAD54645
20	71.4	5.3	135638	7	ABX34289
21	68.6	5.1	88421	6	AAL40781
22	65.4	4.9	484	9	ADE82058
23	65	4.8	897	7	ADA71279

## ALIGNMENTS

## RESULT 1

AAD43514  
ID AAD43514 standard; DNA; 1344 BP.

XX  
AC AAD43514;

XX  
DT 14-NOV-2002 (first entry)

XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #4.

XX  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; ds.

XX  
OS Glycine max.

XX  
FH Key  
FT CDS

FT  
FT Location/Qualifiers  
52..921  
/\*tag= a  
/product= "Maize IPPK protein #4"

XX  
PN WO200259324-A2.

XX  
PD 01-AUG-2002.

XX  
PF 09-JAN-2002; 2002WO-US003120.

XX  
PR 12-JAN-2001; 2001US-0261465P.

XX  
PA (PION-) PIONEER HI-BRED INT INC.

XX  
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;

XX  
DR WPI; 2002-636540/68.

XX  
DR P-PSDB; AAE26196.

XX  
PT New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.

XX  
PS Claim 1; Page 64-65; 86pp; English.

XX  
CC The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or  
CC increasing non-phytate phosphorous to improve the nutritional value of

ADC86736 Human GPC  
AAR25795 S. chryso  
AAL61173 Actinosyn  
AAL61224 Actinosyn  
ABL61294 N. unifor  
AAT93095 Streptomy  
AAV25925 Streptomy  
ADB78935 Rice tran  
ADA71938 Rice gene  
ABS78681 Kitasatos  
AAL17185 Streptomy  
AAL7186 Streptomy  
AAL61181 Actinosyn  
AAV21187 Amycolato  
AAX53491 Human ade  
AAD53058 Streptomy  
AAF74867 Leishmani  
AAX53491 Human ade  
ABZ66716 Orthosomy  
ABZ66808 Orthosomy  
AAD17184 Streptomy  
AAO2484 Human col

CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA  
XX  
SQ Sequence 1344 BP; 228 A; 426 C; 399 G; 291 T; 0 U; 0 Other;

Query Match 100.0%; Score 1344; DB 6; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 1.7e-271;  
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGAGGTCACTGCGTCCATCCCTCGCCCATAGTCCCTTCCCATACCATGTCGAC 60  
DB 1 GACGAGGTCACTGCGTCCATCCCTCGCCCATAGTCCCTTCCCATACCATGTCGAC 60  
QY 61 CTCACCCCGCGAGCACCAAGTCGCGCCGACCGCGCTCCGCCAGCAAGTCGGCCCG 120  
DB 61 CTCACCCCGCGAGCACCAAGTCGCGCCGACCGCGCTCCGCCAGCAAGTCGGCCCG 120  
QY 121 CTCATCGAGGTCCGCGCTCTTCTACAGCGCTCCAGCGCGGACCGTGGGAGCAC 180  
DB 121 CTCATCGAGGTCCGCGCTCTTCTACAGCGCTCCAGCGCGGACCGTGGGAGCAC 180  
QY 181 GAGGTGCGCTTCTATGAGCGGTCTCCGCCAGCGCGCTCCGCCCGCATCCGAGAC 240  
DB 181 GAGGTGCGCTTCTATGAGCGGTCTCCGCCAGCGCGCTCCGCCCGCATCCGAGAC 240  
QY 241 ACCTTCTCCCGGTTCCACGCGACCGACGCTCTCCGCCAGCGCGCGCCCGGGAG 300  
DB 241 ACCTTCTCCCGGTTCCACGCGACCGACGCTCTCCGCCAGCGCGCGCCCGGGAG 300  
QY 301 CCGCATCTCACTGCTGCTCGACGACCTCTCGCGGGGTTTCAGCGCGCTGCGTCCGA 360  
DB 301 CCGCATCTCACTGCTGCTCGACGACCTCTCGCGGGGTTTCAGCGCGCTGCGTCCGA 360  
QY 361 GACATCAAGATCGCGCCATCACTGCGCCACCGAGTTCGCGGAGCGCTTACATCGCCAG 420  
DB 361 GACATCAAGATCGCGCCATCACTGCGCCACCGAGTTCGCGGAGCGCTTACATCGCCAG 420  
QY 421 TGCCTCGCAAGAACCGCGGACACCGAGCTTCTGCTCGATTCGCGCTCCGCGCTC 480  
DB 421 TGCCTCGCAAGAACCGCGGACACCGAGCTTCTGCTCGATTCGCGCTCCGCGCTC 480  
QY 481 CGAGTCTCGCGCCCGAGCGCGCTGTCGCGACCGAGCGCGCGAGGTGAAGCCCATG 540  
DB 481 CGAGTCTCGCGCCCGAGCGCGCTGTCGCGACCGAGCGCGCGAGGTGAAGCCCATG 540  
QY 541 GACACCGCGCGGTCCGCGCGCTGCTCCGCGCTACGTGTCATCCGTTCCGACGAGGGG 600  
DB 541 GACACCGCGCGGTCCGCGCGCTGCTCCGCGCTACGTGTCATCCGTTCCGACGAGGGG 600  
QY 601 ATGGACTGCGCTCCCGCGCGGTGTACGAGGAAAGGTGGAGTCTGTGACAGCTG 660  
DB 601 ATGGACTGCGCTCCCGCGCGGTGTACGAGGAAAGGTGGAGTCTGTGACAGCTG 660  
QY 661 CCGAGCTCAAGCGGTGTCGAGGAGCAGACTCTGTCACCTTCTACTCGCGCTGATT 720  
DB 661 CCGAGCTCAAGCGGTGTCGAGGAGCAGACTCTGTCACCTTCTACTCGCGCTGATT 720  
QY 721 CTTCTGGGTATGATGCTGTCAGTCGAGCGAGGAGATGGGGTGGGTGACGGTG 780  
DB 721 CTTCTGGGTATGATGCTGTCAGTCGAGCGAGGAGATGGGGTGGGTGACGGTG 780  
QY 781 AGCTGGTGAATTTGCCATGTCGAGGAGGTGATGGGTGATTAACCAACTTCCTG 840  
DB 781 AGCTGGTGAATTTGCCATGTCGAGGAGGTGATGGGTGATTAACCAACTTCCTG 840  
QY 841 GCGGGCTCTGCTCGTGTACGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 900  
DB 841 GCGGGCTCTGCTCGTGTACGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 900  
QY 901 CAGCCTTTGGGTCTCTTTAAGAGAGGATCTCTGGCATTTTCGATTTGAACAAAGCCCTA 960

DB 901 CAGCCTTTGGGTCTCTTTAAGAGAGGATCTCTGGCATTTTCGATTTGATTAACAAAGCCCTA 960  
QY 961 CAAGTTTGTCTGGAAGAAAGAGCGCTCCGAGTTGTCTGGGTGTGAGATCTGAGAG 1020  
DB 961 CAAGTTTGTCTGGAAGAAAGAGCGCTCCGAGTTGTCTGGGTGTGAGATCTGAGAG 1020  
QY 1021 GTCGTCGCGCCACTGTTGCTTGCCTTTCGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 1080  
DB 1021 GTCGTCGCGCCACTGTTGCTTGCCTTTCGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 1080  
QY 1081 CTTTTTTTCGCAACCCCTTACTTCCGAAGAAACTTTTTTTTCCCACTTTGGGGGTGGA 1140  
DB 1081 CTTTTTTTCGCAACCCCTTACTTCCGAAGAAACTTTTTTTTCCCACTTTGGGGGTGGA 1140  
QY 1141 TTACGTTGATCTGTTGTTGCTGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 1200  
DB 1141 TTACGTTGATCTGTTGTTGCTGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 1200  
QY 1201 TGATCCGCAACTGTTGCTGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 1260  
DB 1201 TGATCCGCAACTGTTGCTGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 1260  
QY 1261 TCGCATCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
DB 1261 TCGCATCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
QY 1321 TCCGATAAAAAAATAAAAAA 1344  
DB 1321 TCCGATAAAAAAATAAAAAA 1344

RESULT 2  
AAD43513  
ID AAD43513 standard; DNA; 922 BP.  
XX  
AC AAD43513;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #3.  
XX  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; ds.  
XX  
OS Zea mays.  
XX  
FH Key Location/Qualifiers  
FT CDS 53..922  
FT /\*tag= a  
FT /product= "Maize IPPK protein #3"  
XX  
PN WO200259324-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 09-JAN-2002; 2002WO-US003120.  
XX  
PR 12-JAN-2001; 2001US-0261465P.  
XX  
XX (PION-) PIONEER HI-BRED INT INC.  
XX  
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX  
XX WPI; 2002-636540/68.  
XX  
XX P-PSDB; AAE26195.  
XX  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
XX useful in modulating phytic acid biosynthesis by decreasing phytate or  
XX increasing non-phytate phosphorous to improve the nutritional value of  
XX animal feed.  
XX  
PS Claim 1; Page 62-63; 86pp; English.

XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are used to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
XX present sequence is maize IPPK DNA  
SQ Sequence 922 BP; 137 A; 327 C; 288 G; 170 T; 0 U; 0 Other;  
Query Match 64.3%; Score 864.4; DB 6; Length 922;  
Best Local Similarity 98.8%; Pred. No. 38-171;  
Matches 869; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 42 TCCCATACCATGTCGACCTCCACCGCGGAGACCAAGTCGCGGACCGCGCCTC 101  
DB |||||  
QY 102 CCGCAGCAAGCTGGGCGCGCTCATCGAGCGCTCCGCGCTTCTTACAAGCGCTCCAGGC 161  
DB |||||  
QY 162 CGCGACCGTGGGAGCAGAGGTGCGCTTCTATGAGCGTTCTCGGCCACCGCGCGT 221  
DB |||||  
QY 222 CCGCGCGCGATCCGAGACACCTTCTCCCGGTTTCAAGCGACGCGACTCTCCCGCAC 281  
DB |||||  
QY 282 CGAGCGCAGCGCGGAGCGCATCTCACTCGTCTTCAAGCGCTCTCGCGGCTT 341  
DB |||||  
QY 342 TCAGCGCGCTCGTTCGAGACATCAGATCGCGCGCATCAGTGGCCACCGAGTTCGC 401  
DB |||||  
QY 402 GGAGCGCTACATCGCAAGTGTCTCGCAAGCGCGGACCGGACCGAGCGTTCCTCGG 461  
DB |||||  
QY 462 ATTCGCGTCTCCGCGTCCGAGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 521  
DB |||||  
QY 522 CCGGAGTGAAGCGCATGAGACACCGCGCGGTCTCGCGCGGTCTCGCGCGGTACGTGC 581  
DB |||||  
QY 582 ATCCGTTCCGACGAGGAGTGAATGTGCGCTCGCGCGCGGTGTACGAGGAGAAAGG 641  
DB |||||  
QY 642 TGGAGTCTTGTACAGCTGCGCGAGCTCAAGCGGTGGTTCGAGGAGCAGACTCTGTCCA 701  
DB |||||  
QY 702 CTTCTACTCGCGGTGATCTTCTGGCTATGATCTGCTGAGTTCGAGAGCGGAGAG 761  
DB |||||  
QY 762 TGGGGGTGGGTGACGCTGAAGCTGTGACTTTTCCCATGTGGCGGAGGTGATGGGGT 821  
DB |||||  
QY 822 GATTGACCAACAATTCTCTGGGCGGCTCTGCTCGCTGATCAAGTTCGTTTCTGACATGT 881  
DB |||||  
QY 882 GATTGACCAACAATTCTCTGGGCGGCTCTGCTCGCTGATCAAGTTCGTTTCTGACATGT 882

QY 882 TCCGAGAGCTCTCTCATACGAGCGCTTGGTCTTCTTTAA 921  
DB |||||  
QY 922 TCCGAGAGCTCTCTCATACGAGCGCTTGGTCTTCTTTAA 922  
DB |||||  
RESULT 3  
AD43511  
ID AAD43511 standard; DNA; 1169 BP.  
AC AAD43511;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #1.  
XX  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; ds.  
XX  
OS Zea mays.  
XX  
FH Key Location/Qualifiers  
FT CDS 84..806  
FT /\*tag= a  
FT /product= "Maize IPPK protein #1"  
XX  
PN WO200259324-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 09-JAN-2002; 2002WO-US003120.  
XX  
PR 12-JAN-2001; 2001US-0261465P.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX  
DR WPI; 2002-636540/68.  
DR P-PSDB; AAE26193.  
XX  
PT New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX  
PS Claim 1; Page 58-59; 86pp; English.  
XX  
CC The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are used to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA  
SQ Sequence 1169 BP; 238 A; 362 C; 347 G; 222 T; 0 U; 0 Other;  
Query Match 62.9%; Score 844.8; DB 6; Length 1169;  
Best Local Similarity 97.2%; Pred. No. 3.9e-167;  
Matches 890; Conservative 0; Mismatches 22; Indels 4; Gaps 3;  
QY 42 TCCCATACCATGTCGACCTTCCACCGCGGAGACCAAGTCGCGGACCGCGCCTC 101  
DB |||||  
QY 74 TCCCATACCATGTCGACCTTCCACCGCGGAGACCAAGTTCGCGGTGACCGGCTC 133  
DB |||||  
QY 102 CGCAGCAAGCTGGGCGCGTTCATCGAGCGTTCGCGCTTCTTACAAGCGCTCCAGGC 161  
DB |||||  
QY 134 CGCAGCAAGCTGGGCGCGCTTCATCGAGCGTTCGCGCTTCTTACAAGCGCTCCAGGC 193  
DB |||||  
QY 162 CGGCGACCGTGGGAGCAGAGGTGCGCTTCTATGAGCGTTCGCGCCACCGCGCGT 221  
DB |||||

```
Db 194 CGGCGACCGTGGGAGCAGAGTGCCTTCTATGAGGGGTTCTCCGCCACCGCGCGCT 253
QY 222 CCGCGCCCGATCCGAGACACTTCTTCTCCCGGTTCCACGGCAGCGACTCTCTCCCCAC 281
Db 254 CCGCGCCCGATCCGAGACACTTCTTCTCCCGGTTCCACGGCAGCGACTCTCTCCCCAC 313
QY 282 CGAGGCGCAGCCGGGAGCCGATCCTCCTCCTCGTCTCGAGCCTCTCTCGCGGGTT 341
Db 314 CGAGGCGCAGCCGGGAGCCGATCCTCCTCCTCGTCTCGAGCCTCTCTCGCGGGTT 373
QY 342 TCAGGCGCCTCGTGCAGACATCAAGATCGCGCCATCACTGCGCCACCGAGTTCCGCC 401
Db 374 TGAGGCGCCTCGTGCAGACATCAAGATCGCGCCATCACTGCGCCACCGAGTTCCGCC 433
QY 402 GGAGCCCTATCATGCCAAGTGTCTGCCAAGACCGCGGACACAGAGGTTCTCTGTCGG 461
Db 434 GGAGCCCTATCATGCCAAGTGTCTGCCAAGACCGCGGACACAGAGGTTCTCTGTCGG 493
QY 462 ATTCCGCGTCTCGGCGTCCGAGTGTCTGCGGCCCGGAGGCGCGTGTGCGGACCGAGG 521
Db 494 ATTCCGCGTCT--TGCGTCCGAGTGTCTGCGGCCCGGAGGCGCGTGTGCGGACCGAGG 551
QY 522 CCGGAGGTGAAGCCATGACACCGCCCGCGTCTCGCGCGTGTCTCGCGCGTACGTGTC 581
Db 552 CCGGAGGTGAAGCCATGACACCGTCTCGCGTCTCGCGCGTGTCTCGCGCGTACGTGTC 611
QY 582 ATCCG--TTCCGACAGGGGATGATGTGCGCTCTCGCGCGTGTGACGAGGAAAG 640
Db 612 ATCCGTTTCCGACAGGGGATGATGTGCGCTCTCGCGCGTGTGACGAGGAAAG 671
QY 641 GTGGAGTCTGTGACAGCTGCGGAGCTCAAGCGGTGTGACGAGGACAGTCTGTGTC 700
Db 672 GTGGAGTCTGTGACAGCTGCGGAGCTCAAGCGGTGTGACGAGGACAGTCTGTGTC 731
QY 701 ACTTCTACTCGCGTGAATCTTCTGGGCTATGATCTCTGCTGCTGCTGCTGCTGCTGCTG 760
Db 732 ACTTCTACTCGCGTGAATCTTCTGGGCTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 791
QY 761 ATGGGGTGGGTGACGGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 820
Db 792 GTGGGGTGGGTAAACAGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 851
QY 821 TGATTTGACACAACTTCTTGGCGGGCTCTGCTGCTGATCAAGTTCTGATGATGATGATG 880
Db 852 TGATTTGACACAACTTCTTGGCGGGCTCTGCTGATGATGATGATGATGATGATGATGATG 911
QY 881 TTCCGAGACTCTCTATACGACGCTTTGGGTCTTCTTTAAGAGAGGATCTCTGGCA-TTT 939
Db 912 TTCCGAGACTCTCTTAGACGACGCTTTGGGTCTTCTTTAAGAGAGGATCTCTGACATTT 971
QY 940 CGATTTGATAAAG 955
Db 972 TGATTTGATAAAG 987
```

## RESULT 4

AAD43512

ID AAD43512 standard; DNA; 923 BP.

XX AC

AAD43512;

XX AC

DT 14-NOV-2002 (first entry)

XX DT

DE Maize inositol polyphosphate kinase (IPPK) DNA #2.

XX XX

XW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;

XX nutritional value; animal feed; transgenic; gene; ds.

XX OS

XX Zea mays.

XX XX

FH Key Location/Qualifiers

FT 53..736

CDS /\*tag= a

FT FT

```
FT /product= "Maize IPPK protein #2"
XX W0200259324-A2.
XX 01-AUG-2002.
XX 09-JAN-2002; 2002WO-US003120.
XX 12-JAN-2001; 2001US-0261465P.
XX (PION-) PIONEER HI-BRED INT INC.
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX P-PSDB; AAE26194.
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX Claim 1; Page 60-61; 86pp; English.
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are used to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is maize IPPK DNA
XX SQ Sequence 923 BP; 138 A; 326 C; 286 G; 173 T; 0 U; 0 Other;
Query Match 62.3%; Score 837.4; DB 6; Length 923;
Best Local Similarity 97.5%; Pred.No. 1.3e-165; Indels 1; Gaps 1;
Matches 859; Conservative 0; Mismatches 21;
QY 42 TCCCATACCATGTCCGACCTCCACCCCGCGAGCACCAAGTCGCCGCCACCGCGCTC 101
Db 43 TCCCATACCATGTCCGACCTCCACCCCGCGAGCACCAAGTCGCCGCCACCGCGCTC 102
QY 102 GCCAGAGAGCTGGGCGCGCTCATGACGGCTCCGGCTCTTCTACAGAGCGCTCCAGC 161
Db 103 CGCAGCAAGCGCGCGCGCTCATGACGGCTCCGGCTCTTCTACAGAGCGCTCCAGC 162
QY 162 CGGCGACCGTGGGAGCAGAGGTGCGCTTCTATGAGCGCTTCTCCGCCACCGCGCT 221
Db 163 CGGCGACCGTGGGAGCAGAGGTGCGCTTCTATGAGCGCTTCTCCGCCACCGCGCT 222
QY 222 CCGGCGCGCATCCGAGACACTTCTTCCCGCGTTCCACGGCAGCGACTCTCTCCCGC 281
Db 223 CCGGCGCGCATCCGAGACACTTCTTCCCGCGTTCCACGGCAGCGACTCTCTCCCGC 282
QY 282 CGAGCGCAGCCCGGGAGCGCATCTCTCACTGCTCTCGACGACCTCTCTCGCGGGTT 341
Db 283 CGAGCGCAGCCCGGGAGCGCATCTCTCACTGCTCTCGACGACCTCTCTCGCGGGAT 342
QY 342 TCAGGCGCCTTGGTGGCAGACATCAAGATGGGCGCATCACTGCGGCCACCGAGTTCCGC 401
Db 343 TGAGGCGCCTTGGTGGCAGACATCAAGATGGGCGCATCACTGCGGCCACCGAGTTCCGC 402
QY 402 GGAGCCCTATCATGCCAAGTGTCTGCCAAGACCGCGGACCGAGGTTCTCTCGG 461
Db 403 GGAGCCCTATCATGCCAAGTGTCTGCCAAGACCGCGGACCGAGGTTCTCTCGG 462
QY 462 ATTCCGCGTCTCGGCGTCCGAGTGTCTGCGGCCCGGAGGCGCGTGTGCGGACCGAGG 521
Db 463 ATTCCGCGTCTCGGCGTCCGAGTGTCTGCGGCCCGGAGGCGCGTGTGCGGACCGAGG 522
QY 522 CCGGAGGTGAAGCCATGAGCACCGCGCGGTCTCGCGCGTGTCTCGCGCGTACGTGTC 581
```

Db	523	CCCCGAGGTGAAGGCTATGGACACCGTCGGCGCTCCGCGCTGCTCCGCGCTACGTTGTC	582
Qy	582	ATCCG-TTGGCCGACAGAGGGGATGGAGTGTGCTCCTCGCCGCGCGGTGTACGAGGAAGAG	640
Db	583	ATCCGTTTCCCGACAGAGGGGATGGACTCGCCCTCGCGCGCGGTGTACGAGGAAGAG	642
Qy	641	GTGGAGTCTTTGTACAGAGCTGCGCGAGCTCAAGCGTGGTTCAGAGAGCAGACTCTGTTC	700
Db	643	GTGGAGTCTTGTCACTGTGCGCGAGCTCAAGCGTGGTTCGAGGAGCAGCTCTGTTC	702
Qy	701	ACTTCTACTCGCGCTCGATTCTTCTGGGCTATGATGCTGCTCGAGTCGCGAGCAGCGGAG	760
Db	703	ACTTCTACTCGCGCTCGATTCTTCTGGGCTATGATGCTGCTCGAGTCGCGAGCAGCGGAG	762
Qy	761	ATGGGGTGGGGTGACGGTGTAAGCTGTGTGGACTTTTGCCCATGTGGCCGAGGGTGATGGG	820
Db	763	GTGGGGTGGGGTTAACAGTGAAGCTGTGTGGACTTTTGCCCATGTGGCCGAGGGTGATGGG	822
Qy	821	TGATTGACCAACAATTCCTGGCGGGCTCTGCTCGCTGATCAAGTTCTGTTCTGCACATTG	880
Db	823	TGATTGACCAACAATTCCTGGCGGGCTCTGCTAGCTGATCAAGTTCTGTTCTGCACATTG	882
Qy	881	TTCCGGAGACTCCTCATACGACGCGCTTTGGGTCCCTCTTAA	921
Db	883	TTCCAGAGACTCCTCAGACGAGCGCTTTGGGTCCCTCTTAA	923

CC	increasing non-phytate phosphorous to improve the nutritional value of
CC	animal feed, or to reduce the environmental impact of animal waste.
CC	Polynucleotides of the invention are to produce transgenic plants with an
CC	altered phenotype. IPPK proteins are used to screen compounds that
CC	modulate their activity and raising anti-idiotypic antibodies. The
CC	present sequence is maize IPPK DNA
XX	
SQ	Sequence 3416 BP; 893 A; 830 C; 856 G; 837 T; 0 U; 0 Other;
	Query Match 55.1%; Score 740.4; DB 6; Length 3416;
	Best Local Similarity 90.2%; Pred. No. 3.1e-145;
	Matches 862; Conservative 0; Mismatches 31; Indels 63; Gaps 4
QY	2 CACGAGGTCACTCGCTCAACCCCTTCGGCGCCATAGTCCCTTCCCATACCATGTCCGACC 61
DB	28 CACCATCGCCACCGGTCAACCCCTGCTCCATAG-----TCCCATACCATGCCGACC 81
QY	62 TCCACCCGCGGAGACCAAGTCGCGGGCCACCGCGCTTCGCCGACGAAAGTGGGCCCGC 121
DB	82 TCCACCCGCGGAGACCAAGTCGCGGTCAACCGCGCTTCGCCGACGAAAGTGGGCCCGC 141
QY	122 TCATCGAGCGTCCGGCTCTTCTACAGCGCTCCAGCGCGGCGACCGTGGGGAGCAGC 181
DB	142 TCATCGAGGACTCTGGCTCTTCTACAGCGCTCCAGCGCGGCGACCGTGGGGAGCAGC 201
QY	182 AGGTGCGCTTCTATGAGCGGTCTTCGGCCACGCGCGCTCCGGCCCGCATCTCCGAGCA 241
DB	202 AGGTGCGCTTCTATGAGCGGTCTTCGGCCACGCGCGCTCCGGCCCGCATCTCCGAGCA 261
QY	242 CTTTCTTCCCGGTTTCAAGGACGAGACTCTTCCACCGAGCGCAGCCGCGGGAGC 301
DB	262 CTTTCTTCCCGGTTTCAAGGACGAGACTCTTCCACCGAGCGCAGCCGCGGGAGC 321
QY	302 CGCATCTCACTCGTCTCTCGAGCACTCTTCGGGGGTTCAGGCGCGCTGCGTTCGCGAG 361
DB	322 CGCATCTCGACCTCGTCTCTCGAGCACTCTTCGGGGGTTCAGGCGCGCTGCGTTCGCGAG 381
QY	362 ACATCAAGATCGGGCCATCAGTGGCACCGAGTTGCGCGGAGCCCTACATCGCCAAAGT 421
DB	382 ACATCAAGATCGGTGCCATCAGT----- 405
QY	422 GCCTCGCAAGGACCGGGGACCAAGAGCTTCTGCTCGGATTCGCGCTCTCCGCGCTCC 481
DB	406 -----GACCACGAGCGTTCTGCTCGGATTCGCGCTCTCCGCGCTCC 446
QY	482 GAGTCGTGGCCCGAGGGCGCGTGTGCGGACGAGAGCGCCCGAGGTGAAGCCATGG 541
DB	447 GAGTCGTGGCCCGAGGGCGCGTGTGCGGACGAGAGCGCCCGAGGTGAAGGTATGG 506
QY	542 ACACGCGCGCGTCCGCGCGGTGCTCCGGGGCTACGTGTATCCG-TTGCCGACGAGGGG 600
DB	507 ACATTGTGGGTCCGCGCGGTGCTCCGGCGTACGTGTATCCGTTGCCGACGAGGGG 566
QY	601 ATGAGCTGTGGCTTCGCGCGGGGTGTACGAGGAAAAGTGAGTCTTGTTCACAGCTG 660
DB	567 ATGAGCTGTGGCTTCGCGCGGGGTGTACGAGGAAAAGTGAGTCTTGTTCACAGCTG 626
QY	661 CGCGAGCTCAAGGGTGGTTCGAGGACGAGACTCTGTTCACCTTCTACTCGGCGTCAATT 720
DB	627 CGCGAGCTCAAGGGTGGTTCGAGGGGACACTCTGTTCACCTTCTACTCGGCGTCAATT 686
QY	721 CTTCTGGGTATGATGCTGTGAGTCGACGAGCGGAGATGGGGTGGGGTGACGCTG 780
DB	687 CTTCTGGGTATGATGCTGTGAGTCGACGAGCGGAGGTGGGGTGGGGTACAGTG 746
QY	781 AAGCTGATGACTTTTGCCCATGTGGCCGAGGGGTGATGGGGTGAATTGACCAACTTCCTG 840
DB	747 AAGCTGATGACTTTTGCCCATGTGGCCGAGGGGTGATGGGGTGAATTGACCAACTTCCTG 806
QY	841 GCGGGGCTCTGCTGCTGATCAAGTTCGTTCTGACATTTGTTTCGGGAGACTCTCTCAACG 900
DB	807 GCGGGGCTCTGCTGCTGATCAAGTTCGTTCTGACATTTGTTTCGAGAGACTCTCTCAACG 866





```
XX 09-JAN-2002; 2002WO-US003120.
XX 12-JAN-2001; 2001US-0261465P.
XX (PION-) PIONEER HI-BRED INT INC.
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX Claim 1; Page 74; 86pp; English.
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/ or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is maize IPPK DNA
XX
XX Sequence 643 BP; 92 A; 261 C; 178 G; 103 T; 0 U; 9 Other;
XX
XX Query Match 35.0%; Score 470.4; DB 6; Length 643;
XX Best Local Similarity 95.9%; Pred. No. 6.8e-89;
XX Matches 509; Conservative 0; Mismatches 18; Indels 4; Gaps 3;
XX
QY 42 TCCCATACATGTCGACCTCCACCCGCGGAGACCAAGTCGCCGACCGCGCTC 101
DB 107 TCCCATACATGTCGACCTCCACCCGCGGAGACCAAGTCGCCGACCGCGCTC 166
QY 102 CGCCAGCAAGTCGGCGCGCTCATCGACGGTCCGGGCTCTTCTACAGCGCTCCAGGC 161
DB 167 CGCCAGCAAGTCGGCGCGCTCATCGACGGTCCGGGCTCTTCTACAGCGCTCCAGGC 226
QY 162 CGCGGACGTCGGGAGACGACGAGTGGCTTCTATGAGGGTTCGCCGCCACGCGCGGT 221
DB 227 CGCGGACGTCGGGAGACGACGAGTGGCTTCTATGAGGGTTCGCCGCCACGCGCGGT 286
QY 222 CCGGCGCGCATCGAGACACCTTTCTCCCGGTTCCAGGACGCGACTCTCTCCCCAC 281
DB 287 CCGGCGCGCATCGAGACACCTTTCTCCCGGTTCCAGGACGCGACTCTCTCCCCAC 346
QY 282 CGAGGCGAGCCCGGGAGCGCATCTCTCACTCTCTCGACACCTCTCTCGGGGGTT 341
DB 347 CGAGGCGAGCCCGGGAGCGCATCTCTCACTCTCTCGACACCTCTCTCGGGGGTT 406
QY 342 TCAGGCGCGCTCGTCCGACATCAAGATCGGGGCCATCAGTGGGCCACGAGTTGCC 401
DB 407 TGAAGCGCCCTCGTCTCGACATCAAGATCGGGGCCATCAGTGGGCCACGAGTTGCC 466
QY 402 GGAGCCCTACATCGCCAAAGTCCCTCGCCAGGACCGCGGGACCAAGCGTCTCTGTCGG 461
DB 467 GGAGCCCTACATCGNCAAGTACCTNGCCAAAGGACCGCGGGACCAAGCGTCTCTGTCGG 526
QY 462 ATTCGCGCTCTCGGCGTCCGAGTCGTTCGGCCCGGAGGGCGCCGTGTGGCGGACGAGCG 521
DB 527 ATTCGCGCTCT - TGGCGTCCGAGTCGTTCGGCCCGGAGGGCGCCGTGTGGCGGACGAGCG 584
QY 522 -CCGCGAGGTGAAGGCCATGACA - CCGCGCGCGTCTCGCGCGGTCTCTGTCGG 570
DB 585 CCGCGGGGTGAANGCTATGACACCGCTCGGNGNCCGGGNGTGTCTCGG 635
XX
RESULT 8
AAD43520
```

```
ID AAD43520 standard; DNA; 519 BP.
XX
AC AAD43520;
XX
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) DNA #7.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; ds.
XX
OS Zea mays.
XX
FN WO200259324-A2.
XX
PD 01-AUG-2002.
XX
PF 09-JAN-2002; 2002WO-US003120.
XX
PR 12-JAN-2001; 2001US-0261465P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX
PT New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
PS Claim 1; Page 75; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA
XX
SQ Sequence 519 BP; 144 A; 109 C; 126 G; 125 T; 0 U; 15 Other;
```

```
Query Match 22.2%; Score 298.4; DB 6; Length 519;
Best Local Similarity 89.4%; Pred. No. 6.1e-53;
Matches 328; Conservative 0; Mismatches 38; Indels 1; Gaps 1;
QY 627 GTACGGAGGAAAAGGTGGAGTCTTGTACAGCTGCGGAGCTCAAGCGTGGTTTCGAGGA 686
DB 2 GTACGGAGGAAAANGTGGAGTCTTGTACAGCTGCGGAGCTCAANGCGTGGTTTCGAGGG 61
QY 687 GCAGACTCTGTTCCACTTCTACTCGGCGTCGATTCCTTCTGGGCTATGATCTGCTGAGT 746
DB 62 GCAGACTCTGTTCCACTTCTACTCGGCGTCGATTCCTTCTGGGCTATGATCTGCTGAGT 121
QY 747 CGCAGCAGGCGGAGATGGGGTGGGTGACGGTGAAGCTGGTGAACCTTTCGCCATGTGGC 806
DB 122 CGCAGCAGGCGGANGTGGGGGTGGGTGAACAGTGAAGCTGGTGGCTTTCGCCATGTGGC 181
QY 807 CGAGGGTGTATGGGGTGAATGACCAAACTTCTCGGCGGGCTCTGCTCGCTGATCAAGTT 866
DB 182 CGAGGGTGTATGGGGTGAATGACCAAACTTCTCGGCGGGCTCTGCTGATCAAGTT 241
QY 867 CGTTTCTGACATTTTCGGGAGACTCTCTCATACGACGCTTTGGGTCTTCTTTAAGAGAG 926
DB 242 TGTTTCTGACATTTTCCAGAGACTCTCTCAGACGCGAGCTTTGGGTCTTCTTTAAGAAA 301
QY 927 GATCCTGGCA - TTTGATTTGATAA CAAAGCCCTA CAAAGTTTGTCTGGAAGAACGCG 985
DB 302 GATCCTGGCATTTTTCGATTTGATAA CAAAGAACCACTTTTCAGCTGCCAAAAAANCA 361
```

QY 986 CCTCCGA 992  
Db 362 CCACTGA 368

## RESULT 9

AD43521  
ID AAD43521 standard; DNA; 353 BP.

AC AAD43521;

XX 14-NOV-2002 (first entry)

XX Maize inositol polyphosphate kinase (IPPK) DNA #8.

XX Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
XX nutritional value; animal feed; transgenic; ds.

XX Zea mays.

XX WO200259324-A2.

XX 01-AUG-2002.

XX 09-JAN-2002; 2002WO-US003120.

XX 12-JAN-2001; 2001US-0261465P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;

XX WPI; 2002-636540/68.

XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.

XX Claim 1; Page 75; 86pp; English.

XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA

XX Sequence 353 BP; 81 A; 78 C; 91 G; 96 T; 0 U; 7 Other;

Query Match 14.3%; Score 191.6; DB 6; Length 353;  
Best Local Similarity 90.2%; Pred. No. 1.2e-30;  
Matches 238; Conservative 0; Mismatches 21; Indels 5; Gaps 3;

QY 667 CTCAGGCTGCTCGAGGAGAGACTCTGTTCCACTTCTACTCGGCTCGATTCTCTG 726

Db 1 CTCAGGCTGCTCGAGGAGAGACTCTGTTCCACTTCTACTCGGCTCGATTCTCTG 60

QY 727 GGCTATGATGCTGCTGAGTTCGAGCGGAGATGGGGTGGGGTGACGTTGAAGCTG 786

Db 61 GGCTATGATGCTGCTGAGTTCGAGCGGAGATGGGGTGGGGTGACGTTGAAGCTG 120

QY 787 GTGGACTTTGCCATGTGGCCGAGGGTGATGGGT--GATTGACCACAACTTCTTGGGCG 844

Db 121 GTGGACTTTGCCATGTGGCCGAGGGTGATGGGT--GATTGACCACAACTTCTTGGGCG 180

QY 845 GGCTCTGCTGCTGATCAATTC--GTTTCTGACATTTGTCGGAGACTCTCTATCG-C 901

Db 181 AGCTCTGCTAGCTATGATCAATTC--GTTTCTGACATTTGTCGGAGACTCTCTATCGCC 240

QY 902 AGCCTTGGTCTCTTCTTAAGAGA 925  
Db 241 AGCCTTGGTCTCTTCTTAAAAA 264

## RESULT 10

AD43516  
ID AAD43516 standard; DNA; 1195 BP.

XX AAD43516;

XX 14-NOV-2002 (first entry)

XX Eucalyptus grandis inositol polyphosphate kinase (IPPK) DNA.

XX Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;  
KW nutritional value; animal feed; gene; ds.

XX Eucalyptus grandis.

XX Key Location/Qualifiers  
FT CDS 116..1048  
FT /\*tag= a

XX /product= "E. grandis IPPK protein"

XX WO200259324-A2.

XX 01-AUG-2002.

XX 09-JAN-2002; 2002WO-US003120.

XX 12-JAN-2001; 2001US-0261465P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;

XX WPI; 2002-636540/68.

XX P-P5DB; AAE26198.

XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.

XX Claim 1; Page 68-70; 86pp; English.

XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is Eucalyptus grandis IPPK DNA

XX Sequence 1195 BP; 342 A; 240 C; 294 G; 319 T; 0 U; 0 Other;

Query Match 13.7%; Score 184.6; DB 6; Length 1195;  
Best Local Similarity 55.6%; Pred. No. 4.3e-29;  
Matches 471; Conservative 0; Mismatches 349; Indels 27; Gaps 5;

QY 56 CCGACTTCACCCCGGAGCACCAGTGGCGGCGCCCGCTCCGCCAGCAGCTGG 115

Db 114 CCATGCTCAGGTCCCGGATCATCAAGTCGCGGTCAACCGGGAGACGGGGGAAAGCTGG 173

QY 116 GCCCGCTCATCGAGGCTCCGGCTCTTCTACAGCCCTCCAGCGCGCCACCGTGGG 175

Db 174 GCCCAGTGGTGATGATTCGGGCGCGCTCTATAAGCCTCTCCAGAGCATCATCGGAG 233

QY 176 AGCAGAGGTGCGCTTCTATGAGGCGTTCTCCGCCACGCGCGGTCCCGCCCGCATCC 235

```
Db 234 ACACGGAAGTGGCTTTTACAGATCATCTTATTCATATACCGAGATCCAGGTCAATTC 293
Qy 236 GAGACACCTTCTTCCCGGTTCCACGGCAGCGACTCTCTCCACCGAGCGAGCCG 295
Db 294 G---CAAAATCTTCTCGTTCAGCGAACTAAGACTAT--TGAGCGCTCTGATGAT 347
Qy 296 GGGAGCGGCATCTTCACTCTCTCTCGAGCACTCTCTCGCGGTTTCAGCGCCCTGCG 355
Db 348 CGGTCCTCAACCTCACCTGTTCTGGAGATCTCTCTCGGTCGCAAGAACCATCTC 407
Qy 356 TCGCAGACATCAAGATCGCGCCATCAGTCGCGCACCGAGTTCGCGGAGCCCTACATCG 415
Db 408 TCATGACATCAAGCTGGATCCAGACATGATATCGGAGCCCTCTGAGAGTACATCC 467
Qy 416 CCAAGTGCCTCCCAAGACCGCGGACCAAGAGTTCCTCGGATTCGCGCTTCGCG 475
Db 468 AAAAGTGTAGAGAAAGATCGAAATAGCAAGCGTTTCATTCGGTTTATAGGATTTCTG 527
Qy 476 GCGTCGAGTCTGCG---CCCGAGGCGCGGTGTGCGGACCGAGCGCCCGAGGTGA 532
Db 528 GGCTAAGGATATCAAAATAGCGAAGCTGGATTTGGCAACCTGAGAGAAAGTTGTT 587
Qy 533 AGGCATGACACCGCGCGCTCGCGCGCTCGCTCGCGCTACGTGTCATCC----- 585
Db 588 ATAGCTTTAATGCGGACGCTGTCAGTCTGCGCTCTCAGGAAGTTGTTTCTTCAACTGT 647
Qy 586 -----GTTGCGCAGAGGGGATGACTGTGCGCTCGCGCGCGGTGTCAGGAGAAAG 640
Db 648 CTCGGGTCCAAATGTGATCGGATTTGTTGTATGTCATCAAAAGTTTACTGTCAACGGG 707
Qy 641 GTGAGTCTTGTACAGCTGCGGAGCTCAAGGCGTGGTTCGAGGAGCAGACTCTGTTC 700
Db 708 GTGGAATTTGGCAAAATGCTTACGTCGAAGGATGTTTGGAGTTCAGACGAATATC 767
Qy 701 ACTTCTACTCGCGTCGATTTCTTGGGCTATGATGCTGTGTCAGTCGACAGCGCGAG 760
Db 768 ACTTCTATCTTGTCTACTCATTTATTTATATGACAGGAGTCTGCTTTGACGGC--- 823
Qy 761 ATGGGGTGGGGTGACGGTGAAGCTGTGGGACTTTTGGCCATGTGGCCGAGGATGGG 820
Db 824 --TGTGACACACCCGAAAGTAAACTGTGTGACTTTTGCACATGTGATGATGCGCCACGCG 881
Qy 821 TGATTGACCAACTTCTCGGGCGGCTCTCTCGCTGATCAAGTTCGTTTCTGACATTG 880
Db 882 TGATGATCACAACCTTCTGGGTGGCTCTGTTCTGTAATCAAGTTTATACGTGACATG 941
Qy 881 TTCGGGA 887
Db 942 CTGATGA 948
```

## RESULT 11

```
AAD43515
ID AAD43515 standard; DNA; 1105 BP.
XX AC
AC AAD43515;
XX
DT 14-NOV-2002 (first entry)
XX
DE Soybean inositol polyphosphate kinase (IPPK) DNA.
XX
KW Soybean; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
CDS 12..851
FT /*tag= a
FT /product= "soybean IPPK protein"
XX
PN WO200259324-A2.
```

```
XX 01-AUG-2002.
XX 09-JAN-2002; 2002WO-US003120.
XX 12-JAN-2001; 2001US-0261465P.
XX (PION-) PIONEER HI-BRED INT INC.
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX P-PSDB; AAE26197.
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX Claim 1; Page 66-67; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is soybean IPPK DNA
XX
XX Sequence 1105 BP; 279 A; 311 C; 233 G; 282 T; 0 U; 0 Other;
XX
XX Query Match 13.3%; Score 179; DB 6; Length 1105;
XX Best Local Similarity 55.6%; Pred. No. 6.3e-28;
XX Matches 460; Conservative 0; Mismatches 340; Indels 27; Gaps 5;
Qy 61 CTCACCGCGGAGCACCAGTCGCGCGCACCGCGCTCCGACAGCTGGGCGC 120
Db 15 CTCAAGATCCCGAGCACCAGGTGGCGGACAAAGCCCAAGGACCGAATCTTGGGCCA 74
Qy 121 CTCATCAGCGCTCCGGCTCTTCTACAAGCGCTTCCAGGCGCGCGACCGTGGGAGCAC 180
Db 75 CTCGTACAGATTTGGAAATCTACAAGCCCTCCAGACCAACAAGACGACGACACC 134
Qy 181 GAGTCTCGCTTCTATAGAGCGGT-----CTCGCCCAACGCGCGCTCCCGCGCGATC 234
Db 135 CGCGCTCCACCGAACTCTCTTTTACACCTCTCTCGCGCGCGCGCCACGACTACTCC 194
Qy 235 CGAGACACCTTCTTCCCGGTTCCAGGACGCGACTCTCTCCCGACGAGGCGCAGGCC 294
Db 195 ATCGCTCTCTTCTTCCCGGCTTTCAGGACACCGCTCTCTGAGCGCTCCGACGGTCC 254
Qy 295 GGGAGCGCGATCTCTACCTCTCTCGACGACCTCTCTCGCGGGTTTCAGCGCGCTTCG 354
Db 255 GG---TCCCGACCTCTCTCTGAGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 311
Qy 355 GTCGACACATCAGATCGGGGCTATCAGTGGCACCGAGTTCGCGGAGCCCTACATC 414
Db 312 GTCATGGACGTAAGATCGGCTCCAGAACCTGGGAGACTCCGAGGACTACATC 371
Qy 415 GCCAAGTCTTCGCAAGGACCGCGGACCAAGCGGTCTCTGCTCGGATTCGGGTCTCC 474
Db 372 TGAAGTCTCTGAAGAGGACAGAGATCTCTTAGTTCCTCTGCTTGGGTTTCAGATCTCG 431
Qy 475 GGCCTCGAGTCTGTCGGCGCGCGGCGGTGTGTGGGAGCGGAGCGCCCGAGGTGAAG 534
Db 432 GGACTCAAGACTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 491
Qy 535 GCCATGACACCGCGCGGTCTCGCGCGGTCTCGCGCGCT---ACGTGTCTATCTCGTGGC 591
Db 492 GCCCATGCTGTGCACTTGTCTCTCAAGAGTTCGTTTCTCTAATAATATCAACCATGAT 551
```

QY 592 GACGAGGGATGGAAGTGTGCTGCGCGCGGTGTACGGAGGAAAAGGTGGAGTCTTG 651  
Db |||||  
552 GATCATATCCGATTCGGCTTCGCAACGGAGGTCTAC-----GGGCGGTG 602  
QY 652 TCACAGCTGCGGAGCTCAAGCGTGTTCGAGGAGAGACTCTGTTCCACTTCTACTCG 711  
Db |||||  
603 GAGCGCTGCAAGAGCTCAAGAGCTGGTTCAGAGTTCAGACGGGTATCACTTCTATTCT 662  
QY 712 GCCTCGATTCTCTGGGCTATGATCTGCTGCTGCGAGCGAGGAGATGGGGTGG 771  
Db |||||  
663 TGTTCGTCTCTGTGTGTACAGAGGATCTAGGAAA-----GGAAAGCTTACCAAC 716  
QY 772 GTGAGGTGAAGCTGTGTGACTTTGCGCATGTGCGCGAGGTGATGGGGTGAATGACAC 831  
Db |||||  
717 CCTGTGTCAAACCTCGTGTGACTTTGCAACGTGTGGACGGAAACGGGTGATGATCAC 776  
QY 832 AACTTCTCTGGCGCGGTCTGCTGCTGATCAAGTTCTGATGACAT 878  
Db |||||  
777 AACTTCTGGGTGGCCCTTGTCTCTCATCAAGTTCTCAAGGATAT 823

## RESULT 12

AAD43517

ID AAD43517 standard; DNA; 1020 BP.

XX AC

XX AAD43517;

XX AC

XX DT 14-NOV-2002 (first entry)

XX AC

XX DE P. argentatum inositol polyphosphate kinase (IPPK) DNA #1.

XX KW

XX EN Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;

XX KW nutritional value; animal feed; gene; ds.

XX KW

XX OS Parthenium argentatum.

XX FH

XX Key Location/Qualifiers

XX CDS 21..908

XX FT /\*tag= a

XX FT /product= "P. argentatum IPPK protein #1"

XX FT

XX PN W0200259324-A2.

XX XX

XX PD 01-AUG-2002.

XX XX

XX PF 09-JAN-2002; 2002WO-US003120.

XX XX

XX PR 12-JAN-2001; 2001US-0261465P.

XX XX

XX PA (PION-) PIONEER HI-BRED INT INC.

XX XX

XX PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;

XX XX

XX DR WPI; 2002-636540/68.

XX DR

XX P-PSDB; AAE26199.

XX XX

XX PT New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.

XX PT

XX PS Claim 1; Page 71-72; 86pp; English.

XX XX

XX CC The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is P. argentatum IPPK DNA

XX CC

SQ Sequence 1020 BP; 297 A; 195 C; 219 G; 309 T; 0 U; 0 Other;

Query Match 11.1%; Score 149.8; DB 6; Length 1020;  
Best Local Similarity 53.2%; Pred. No. 8e-22;  
Matches 456; Conservative 0; Mismatches 362; Indels 39; Gaps 5;

QY 61 CTCACCCGCGGAGCACCAGTCCGCGCCACCGCGCTCCGCGACAGAGTGGGCGG 120  
Db |||||  
24 CTCAGGCCCCCAGATCATCAGTTGTGACATGAAGCTGGGCTGGGAAGCTTGGCCCA 83  
QY 121 CTCATGACGGCTCCGCGCTCTTCTACAAGCCGCTCCAGGCCGCGACCGTGGGGAGCAC 180  
Db |||||  
84 CTCATTGATGATTGAGGCCGGTTTACAAACCACTGCAGGGTGATTAACCGTGGGTCA 143  
QY 181 GAGGTGCGCTTCTATGAGGGGTCTCCGCCACGCGCGCTCCGCGCCGCGATCCGAGAC 240  
Db |||||  
144 GAAGTAGCCCTTTATGAATCAITTTCTTAAATAATAATTAATTAATTAATTAATTA 200  
QY 241 ACCTTCTTCCCGGTTTCCAGCGCAGCTCCTCCACCGAGGCGAGCGCGCGGAG 300  
Db |||||  
201 AAATTTCTTCTATATATATATATGSCACCAAAAT-----CATGAAGGCATCCACTGGCTCT 254  
QY 301 CCGCATCTCTCAGCTGCTCTCGAGACCTCTCCGCGGGGTTTCAGGGCGCTCGCTCGCA 360  
Db |||||  
255 GACCATCTCTCAGATGGTGTGCAAGATCTTACATCAGCTCATGTCAACCCATCTGTAATG 314  
QY 361 GACATCAAGATCGCGCCCATCAGTGGCCACCGAGTTCCGCGGAGCCCTACATCGCCAAG 420  
Db |||||  
315 GACATCAAAATCGGGTCCAGAACATGGCGCCAGAAAGCTTCCGAGGCGTACATTCGCAAAA 374  
QY 421 TGCTTCGCCCAAGGACCGCGGACACGAGGGTCTCTGTGCGATTCCGCGCTCCCGCGTC 480  
Db |||||  
375 TGCTTAAAAAGGATAGGAAAGCAAGTATTCATTCGGAATTCAGGATCTCCGGGCTG 434  
QY 481 CGAGTCTCGGCCCGAGGGCGCGTGGCGGACGG---AGCGCCGCGAGGTGAAGGCC 537  
Db |||||  
435 CAACTCTATATCGATGATGGTTCAGGGTTTATAGCCTCATAGAAATTACATGCGTAAA 494  
QY 538 ATGGACACCGCGCGCTCCGCGCGTCCGCGGTAGTGTCTATCCCTGCGGAGCGAG 597  
Db |||||  
495 ACCGCGCCAGCTGATGTTAGACTTCTTAGAAAATTTGTTTCTTAACCCGTCGCA 554  
QY 598 GGGAT-----GGACTGTGCGCTCGCCGCGGCGGTACGGA 633  
Db |||||  
555 GAGATGGAATGCGCACAGCGCTAGCCCGGATTGTTCTTTAGCATCTTTGTTATGGT 614  
QY 634 GGAAGAGTGGAGTCTTGTCAAGTGGCGGAGCTCAAGCGGTGGTTGAGGAGAGACT 693  
Db |||||  
615 GGGCCTAATGGGATATAGCTCAACTGATGGAATGAAGACATGGTTTGAAGATCAACA 674  
QY 694 CTGTTCCACTTCTACTCGGCGTCAATCTTCTGGGCTATGATGCTGCTCAGTCGAGCA 753  
Db |||||  
675 ATTACCACTTCTATGCTGTGTTCTTTTGTTCATCTTTGAAAGAGGTGGTGTGTTAAA 734  
QY 754 GCGGAGATGGGGTGGGTGACGGTGAAGTGGTGAATTTGCCCATGTGCCCGAGGGT 813  
Db |||||  
735 GGT---GCTCGGTCAAAACGAGAAGTCAAACTTATGATTTGCTCATGTTACAGATGT 791  
QY 814 GATGGGTGATTGACACAACTCTTGGCGGGCTCTGCTCGCTGATCAAGTTCGTTCT 873  
Db |||||  
792 AATGGTGTATTGATCACAAAATTTCTTGGTGGGCTCTGTTCTTGTATGAAGTTCTATTTCT 851  
QY 874 GACATTGTTCCGGAGAC 890  
Db |||||  
852 GACATACCTTCGGAGAC 868

## RESULT 13

AAC39023

ID AAC39023 standard; DNA; 1168 BP.

XX AC

XX AAC39023;

XX XX



PR	27-AUG-1999;	99US-0151066P.	
PR	27-AUG-1999;	99US-0151080P.	
PR	30-AUG-1999;	99US-0151303P.	
PR	31-AUG-1999;	99US-0151438P.	
PR	01-SEP-1999;	99US-0151930P.	
PR	07-SEP-1999;	99US-0152363P.	
PR	10-SEP-1999;	99US-0153707P.	
PR	13-SEP-1999;	99US-0153758P.	
PR	15-SEP-1999;	99US-0154018P.	
PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154779P.	
PR	22-SEP-1999;	99US-0155139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	29-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157753P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	14-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	26-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	28-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	

Query Match 11.1%; Score 148.6; DB 3; Length 1169;  
Best Local Similarity 53.0%; Pred.No. 1.5e-21;  
Matches 453; Conservative 0; Mismatches 369; Indels 33; Gaps 5;

QY	49	ACCATGTCGACCTCCACCGCGGAGACCAAGTCGCGGCGCACCGGCGCTCCGCCAGC 108
DB	170	AACAGATGCAGCTCAAGTCCCTGACATCAGTTGAGGACACATTGCTAAACACGGG 229
QY	109	AGCTGGGCGCCCTCATCGACGGCTCCGCCCTCTTCTACAGCGCTCCAGCGCGGAC 168
DB	230	AAGCCTGGTCCCTCTCGTAGATGACAGGGTCGGTCTTCAAGCCACTTCAGGGCGATTCT 289
QY	169	CGTGGGACACAGAGTCCCTTCTATGAGGCTTCTCGGCCACCGCGCTCCGCCGCC 228
DB	290	CGTGGTAAATCGAGTAAAGTCTACGAATCTTCTCTCTCAACACACAGAGTCCAGAA 349
QY	229	CGCATCCGAGACACCTTCTTCCCGGTTCCACGGCAGCGACTCTCTCCACCGAGCGG 288
DB	350	CACATCC---ATAGATATTTCCGGTGTATCAGGCACCTCA-----GCATTT 394
QY	289	CAGCCCGGAGCGCGCATCCTCACTCGCTTCGACGACCTCTCGCGGGGTTTCAGGCG 348

DB	395	GAAGGTTCTGATGGAGGACGCCATGATGGTGTGGAAAAATCTTCTTCGAGATACTCAAA 454
QY	349	CCCTGGTCGAGACATCAAGATCGCGGCCATCACTGTCGCCACCGAGTTTCGCGGAGCCC 408
DB	455	CCATCAGTAATGGATGTTAAGATGGGTTCCAGAACATGGTATCCTGATGCATCTGAAGAA 514
QY	409	TACATCGCCCAAGTCTCGCCCAAGGACCGCGGAGCACAGAGGTTCTGCTCGATTCGCG 468
DB	515	TACATCCAAAATGTTTGAAGAAAGACACGGGTACCAACACCGTGTCTATCGGGTTTCAGG 574
QY	469	GTCTCGGCGTCCGAGTCTGTCGGCCCGGAGGCGGC---GTGTGGCGGACGAGCGCCCG 525
DB	575	ATCTCTGGTTTCGAAGTGTATGATCACAAAGATCGAGTTTCTGGAAGCCCGAGGGAAG 634
QY	526	GAGGTGAAGGCCATGGACACCGCCCGGTCGCCCGCTCCGCGCTTACGTGCA--- 582
DB	635	CTTCTTCGCGGCTCGATGTAGATGAGCGAGATTGACTCTGAGGAAGTTTGTATCATCT 694
QY	583	-----TCCGTTGCGGACGAGGGGATGGACTGTGCGCTCGCGCGCGGCTGTACGGA 633
DB	695	AACTCACTTTCGACACTGGCTCGAAACCTGACTCTGCTTTTGCCTCGAGTGTTCACGGC 754
QY	634	GGAAAAGGTGGAGTCTTGTTCACAGCTGCGCGAGCTCAAGGCGTGGTTTCAGGAGCAGACT 693
DB	755	GGTTCGCCACGGATCTTAACGCGAGTTGCTGGAACCTCAAGACCTGGTTCGAGAACCAAG 814
QY	694	CTGTTCCACTTCTACTCGGCGTCTGATCTTCTTGGGCTATGATGCTGCTGCAATC---GCA 750
DB	815	CTTACCATTTCAACTCTTGTTCGATTTTAAATGGTCTATGAGATGAATCCATCTTGAAG 874
QY	751	GCAGCGGAGATGGGGTGGGTTGACGGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 810
DB	875	GGAAATGATGATGATGCTAGACCAAGTCAAGCTGGTGGATTTTGCCTCATGTTCTTGAT 934
QY	811	GGTATGGGTTGATGACCAAACTTCTGCGCGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 870
DB	935	GGTATGGTGTGATGACCAAACTTCTTGGGTGGTCTTGTCTTCTTCTTCTTCTTCTTCTT 994
QY	871	TCTGACATGTTCCG 885
DB	995	CGTGAGATTTCTCAG 1009

RESULT 14  
AAC33685  
ID AAC33685 standard; DNA; 1104 BP.  
XX  
AC AAC33685;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 3949.  
XX  
KW Hybridisation assay; Genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.





PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match			10.9%; Score 147; DB 3; Length 1104;
Best Local Similarity			52.9%; Pred. No. 3.1e-21;
Matches 452; Conservative 0; Mismatches 370; Indels 33; Gaps 5;			
QY	49	ACCATGTCGACCTCCACCCCGGAGCACCAAGTCGGCGGCACCGCGCTCCGCGAGC	108
DB	120	AACAAGATGACGCTCAAGTCCCTGAACATCAGGTTGCAGGACACATTCGTTAAAGACGGG	179
QY	109	AAGCTGGGCGCGCTCATGACGGCTCCGGCGCTCTTCTACAGCGCTCCAGCGCGCGGAC	169
DB	180	AAGCTGCTCTCTCTAGATGACAAGGTCGGTCTTCAAGCCACTTCAGGGCGATTCT	239
QY	169	GTGGGAGCAGGAGTCGCGCTTCTATGAGCGTCTCCGCCACCGCGCGCTCCCGGC	228
DB	240	CGTGGTGAATCGAGGTAAGTTCAGGAATCTTCTCTCAACACAGAGGTTCCAGAA	299
QY	229	CGATCCGACACACTTCTTCCCGGTTCCAGCGCAGCGACTCTCCCGACCGAGCG	288
DB	300	CACATCC---ATGATATTTCCCGGTATACGGCACTCAA-----GCAGTT	344
QY	289	CAGCCCGGAGCGCATCTCTACCTCGCTCGACACTCTCGCGGGTTTCAGCG	348
DB	345	GAAGGTTCTGATGAGCAGCCATGATGTTGTTGGAATCTTTCAGAAATCTCAAAA	404
QY	349	CCCTGGTCGACACATCAAGATCGGCGCCATCAGTGGCCACCGAGTTTCGCGGAGCC	408
DB	405	CCATCAGTAATGATGTTAAGATGGTTCGAGAACATGGTATCTGATGATCTGAAGA	464
QY	409	TACATCCGCAAGTCTCGCCAGGACCGCGGACCAAGCGCTTCTCGGAGTTCCG	468
DB	465	TACATCAAAATGTTTGAAGAAGACACGGGTACCAACACCGTGTATCGGGTTTCAGG	524
QY	469	GTCTCGGCGTCCGAGTCTCGCGCCCGGAGGGCGCC---GTGGGGGACGAGCGCCG	525
DB	525	ATCTCTGTTTCAAGATGATGATCAAGAATCGAGTTTCGGAAGCCCGAGAGAA	584
QY	526	GAGTGAAGCCATGACACCGCGCGCTCCCGCGCTCCCGCGCTACGTGCA---	582
DB	585	CTTCTTCGCGGCTCATGTAGATGAGCGAGATTGACTCTGAGGAAGTTTGTATCATCT	644
QY	583	-----TCGTTCCGACGAGGGATGGACTGTGCGCTCCCGCGCGGTGTACGGA	633
DB	645	AACACTCTTCGACACTGGCTCGAAACCTGCTCTGCTTTTGCCTCGAGTGTTCAGGC	704
QY	634	GGAAGAGTGGAGTCTTGTACAGCTTCGCGAGCTCAAGCGTGGTTCGAGGACGACT	693
Db	705	GGTCCACGGGATCTTAACGGAGTTGCTGGAACTCAAGACCTGGTTCGAGAACCAACG	764
QY	694	CTGTTCCACTTCTACTCGGCGTCGATTCTTCTGGGCTATGATGCTGCTCAGTC---GCA	750
DB	765	CTCTACCAATTCACACTCTTGTTCGATTTAATGGTCTATGAGAATGAATCCATCTTGAAG	824
QY	751	GCAGCGGAGATGGGGGTGAGCGGTGAAGCTGGTGGACTTTGCCCATGTGCGCGAG	810
DB	825	GGAAATGATGATGATGCTAGACCAACAGTCAAGCTGGTGGATTTCATCTTCTTGAT	884
QY	811	GGTATGGGTGATTGACCAACACTTCTCGGCGGCTCTGCTGCTGATCAAGTTCGTT	870
DB	885	GGTAATGGTGCATTGACCAATACTTCTGGGTGCTTTGCTTTTATAAACTTCATT	944
QY	871	TCTGACATGTTCCG	885
DB	945	CGTGAGATTCCTCAG	959
RESULT 15			
AAC38692			
ID	AAC38692 standard; DNA; 1243 BP.		
XX	AC AAC38692;		
XX	17-OCT-2000 (first entry)		
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 21894.		
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
OS	Arabidopsis thaliana.		
XX	EP1033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-00301439.		
PR	25-FEB-1999;	99US-0121825P.	
PR	05-MAR-1999;	99US-0123180P.	
PR	09-MAR-1999;	99US-0123548P.	
PR	23-MAR-1999;	99US-0125788P.	
PR	25-MAR-1999;	99US-0126264P.	
PR	29-MAR-1999;	99US-0126785P.	
PR	01-APR-1999;	99US-0127462P.	
PR	06-APR-1999;	99US-0128234P.	
PR	08-APR-1999;	99US-0128714P.	
PR	19-APR-1999;	99US-0129845P.	
PR	21-APR-1999;	99US-0130077P.	
PR	23-APR-1999;	99US-0130449P.	
PR	23-APR-1999;	99US-0130510P.	
PR	28-APR-1999;	99US-0130891P.	
PR	30-APR-1999;	99US-0131449P.	
PR	30-APR-1999;	99US-0132048P.	
PR	04-MAY-1999;	99US-0132407P.	
PR	05-MAY-1999;	99US-0132484P.	
PR	06-MAY-1999;	99US-0132485P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	07-MAY-1999;	99US-0132487P.	
PR	11-MAY-1999;	99US-0132863P.	
PR	14-MAY-1999;	99US-0134256P.	
PR	14-MAY-1999;	99US-0134218P.	
PR	14-MAY-1999;	99US-0134219P.	
PR	14-MAY-1999;	99US-0134221P.	
PR	18-MAY-1999;	99US-0134370P.	
PR	19-MAY-1999;	99US-0134768P.	
PR	20-MAY-1999;	99US-0134941P.	
PR	21-MAY-1999;	99US-0135124P.	
PR	24-MAY-1999;	99US-0135353P.	
PR	24-MAY-1999;	99US-0135629P.	

PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137522P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 18-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140895P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141087P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 06-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 12-JUL-1999; 99US-0142970P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 27-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 11-AUG-1999; 99US-0148171P.  
PR 12-AUG-1999; 99US-0148313P.  
PR 13-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0149910P.  
PR 26-AUG-1999; 99US-0150566P.  
PR 27-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151130P.  
PR 01-SEP-1999; 99US-0151438P.  
PR 07-SEP-1999; 99US-0152330P.  
PR 10-SEP-1999; 99US-0152363P.  
PR 13-SEP-1999; 99US-0153070P.  
PR 15-SEP-1999; 99US-0153758P.  
PR 16-SEP-1999; 99US-0154018P.  
PR 20-SEP-1999; 99US-0154039P.  
PR 22-SEP-1999; 99US-0154779P.  
PR 23-SEP-1999; 99US-0155139P.  
PR 24-SEP-1999; 99US-0155486P.  
PR 28-SEP-1999; 99US-0155659P.  
PR 29-SEP-1999; 99US-0156458P.  
PR 04-OCT-1999; 99US-0156596P.  
PR 05-OCT-1999; 99US-0157117P.  
PR 06-OCT-1999; 99US-0157753P.  
PR 07-OCT-1999; 99US-0157865P.  
PR 08-OCT-1999; 99US-0158029P.  
PR 12-OCT-1999; 99US-0158232P.  
PR 13-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 21-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 22-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 25-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 26-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 28-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.

PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

### Query Match

Query Match 10.9%; Score 147; DB 3; Length 1243;  
Best Local Similarity 52.9%; Pred. No. 3.2e-21;  
Matches 452. Conservation

Matches	452; Conservative	0; Mismatches	370; Indels	33; Gaps	5;
QY	49	ACCATGTCGACCTCCACCCGCGGAGACACCAAGTCGCGGCACCGCGCCTCGCCAGC	108		
Db	257	AACAAGATGCAGCTCAAAAGTCCTGAAATCATCAGGTGCGAGACACATGCTAAAGACGGG	316		
QY	109	AAGCTGGGCCCTCATCGACGGCTCCGGCCTCTTCTACAAGCCGCTCCAGGCGCGGCAC	168		
Db	317	AAAGCCTGGTCCTCTCGTAGATGACAAAGGTGGTCTTTCAGCCACCTTCAGGGCGAATCT	376		
QY	169	CGTGGGAGCAGAGGTGCGCTTCTATGAGCGGTTCTCGGCCACGCGCGCTCCCGGCC	228		
Db	377	CGTGTGAAATCGAGGTAAAGTCTACGAATCTTCTCTCTCAAAACAGAGGTTCCAGAA	436		
QY	229	CGCATCCGAGACACCTTCTTCCCGGGTTCACGCGCAGCGACTCTCCCCACGAGGCG	288		
Db	437	CACATCC---ATAGATATTTCCCGGTGTATCAGCGCACTCAA-----CGATT	481		
QY	289	CAGCCCGGAGCCGCATCTCACTCGTCTCGACGACCTCTCGCGGGTTTCAGGCG	348		
Db	482	GAAGGTTCTGATGGAGCAGCCATGATGGTGTGAAAAATCTTCTTCAGATACTCAAA	541		
QY	349	CCCTGCGCTCGCAGACATCAAGATCGGCGCCATCACGTGGCCACCGAGTTCCGCGAGCC	408		
Db	542	CCATCAGTAATGGATGTTAAGATGGGTTGAGAACATGCTATCTGTGATCATCTGAAGAA	601		
QY	409	TACATCGCCAAAGTCCTCGCCAAAGGACCGCGGACCAACGAGCGCTTCTGCTCGGATCCGC	468		
Db	602	TACATCAAAAAATGTTTGAAGAAGACACGGGTACCAACAACCGTGTATCGGGTTTCAGG	661		
QY	469	GTCTCCCGCGCTCCGAGTGTGTCGCGCCCGAGGGCGCC---GTGTGGCGGACGGAGCGCCG	525		
Db	662	ATCTCTGGTTTTCGAAGTGTAATGATCAAAAGATCGAGTTTCTTGAAGCCCGAGAGGAAG	721		
QY	526	GAGGTGAAGGCCATGACACCGCGCGCTCCGCGGTGCTCCGGCGTACGTGTCA---582			
Db	722	CTTCTTCGCGGGCTCGATGTAGATGAGCGAGATTGACTCTGAGGAAGTTTGTATCATCT	781		
QY	583	-----TCCGTTGCCGACGAGGGGATGGACTGTGCTCGCCGCGCGCGGTGTACGGA	633		
Db	782	AACTCACTTTTCGACACTGGCTCGAAACCTGACTCTGCTTTTGGCTCGAGTGTTCACGCG	841		
QY	634	GGAAAAGGTGGAGTCTTGTACAGCTGCGGAGCTCAAGCGGTGGTTCGAGGAGCAGACT	693		
Db	842	GGTTCACCGGATCTTAACGAGTTGCTGSAACTCAAGACCTGGTTCGAGAACCAACG	901		
QY	694	CTGTCCACTTCTACTCGGCGTCGATTCTTCTGGGCTATGATGCTGTGAGTC---GCA	750		
Db	902	CTTACCAATTCAACTCTTTGTTCCGATTTTAATGGTCTATGAGATGAATCAATCTGAAG	961		
QY	751	GCAGGCGGAGATGGGGTGGGTGACGGTGAAGCTGGTGAATTTGGCCATGTGGCGGAG	810		
Db	962	GGAAATGATGATGCTAGACCAAGAGTCAAGCTGGTGGATTTTGTGCTCATGTTCTTGAT	1021		
QY	811	GGTGATGGGGTGATTAACACAACTTCTCGGCGGGCTCTGCTCGCTGATCAAGTTGTT	870		
Db	1022	GGTAATGGTGTCAATGACCAATCACTTTGGGTGGTCTTGTCTCTTCAATAAATTCATT	1081		
QY	871	TCAGCAATGTTCCG	885		
Db	1082	CGTGAGATCTTCAG	1096		

Search completed: March 27, 2004, 02:03:51  
Job time : 610 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2004, 01:44:37 ; Search time 118 Seconds  
(without alignments)  
6320.797 Million cell updates/sec

Title: US-10-042-894A-7  
Perfect score: 1344  
Sequence: 1 gcacgaggtcagtcgctcac.....ataaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq:  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq:  
5: /cgn2\_6/prodata/2/ina/PCRTUS\_COMB.seq:  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69.6	5.2	1926	4	US-09-249-585A-4
2	69.6	5.2	1931	2	US-09-130-114-2
3	58.6	4.4	38506	3	US-09-320-878-19
4	58.6	4.4	38506	4	US-09-141-908-1
5	58.6	4.4	38506	4	US-09-657-440-19
6	58.6	4.4	4411529	3	US-09-103-840A-1
7	58.6	4.3	1082	4	US-09-881-165-4
8	57.6	4.3	23673	4	US-09-773-816-1
9	57	4.2	11220	3	US-09-105-537-32
10	57	4.2	36778	3	US-09-105-537-5
11	55.2	4.1	15872	3	US-09-105-537-1
12	55.2	4.1	15872	4	US-09-091-609-1
13	55.2	4.1	15872	4	US-09-091-609-3
14	55.2	4.1	4403765	3	US-09-103-840A-2
15	54.8	4.1	1704	1	US-08-528-199-2
16	54.8	4.1	1704	1	US-08-528-199-5
17	54.8	4.1	4403765	3	US-09-103-840A-2
18	54.4	4.0	985	4	US-09-056-556-182
19	54.4	4.0	985	4	US-09-072-596-177
20	54.4	4.0	985	4	US-09-072-596-177
21	54.4	4.0	4411529	3	US-09-103-840A-1
22	53.8	4.0	1333	4	US-09-372-422A-9
23	53.4	4.0	921	4	US-09-252-991A-6598
24	53.4	4.0	3084	4	US-09-252-991A-6639
25	53.4	4.0	3273	4	US-09-252-991A-6578
26	53.2	4.0	6854	4	US-09-194-905-7
27	53.2	4.0	152331	3	US-09-128-155-16
					Sequence 4, Appli
					Sequence 2, Appli
					Sequence 19, Appli
					Sequence 19, Appli
					Sequence 1, Appli
					Sequence 4, Appli
					Sequence 1, Appli
					Sequence 32, Appli
					Sequence 5, Appli
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 3, Appli
					Sequence 2, Appli
					Sequence 5, Appli
					Sequence 2, Appli
					Sequence 182, App
					Sequence 177, App
					Sequence 182, App
					Sequence 1, Appli
					Sequence 9, Appli
					Sequence 6598, Ap
					Sequence 6639, Ap
					Sequence 6578, Ap
					Sequence 7, Appli
					Sequence 16, Appli

## RESULT 1

US-09-249-585A-4

; Sequence 4, Application US/09249585A

; Patent No. 6417002

; GENERAL INFORMATION:

; APPLICANT: Horlick, Robert

; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES

; FILE REFERENCE: 0867/05905

; CURRENT APPLICATION NUMBER: US/09/249,585A

; CURRENT FILING DATE: 1999-02-11

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4

; LENGTH: 1926

; TYPE: DNA

; ORGANISM: Epstein Barr Virus

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)-(1926)

; OTHER INFORMATION: template strand of EBNA-1 DNA

US-09-249-585A-4

Query Match

Best Local Similarity 5.2%; Score 69.6; DB 4; Length 1926;

Matches 246; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY	39	CCITCCCGATACATGTCGGACCTCCACCCGCGGAGCACCAAGTCGCGCGCCACCGGCGC	98
Db	274	CCTCGTCTCTCGCCCTCCCGCT	333
QY	99	CTCCGCGCAGCAAGTCGGCCCGGTCTCGAGGGTCTCGGCTCTCTCTCTCTCTCTCTCTCT	158
Db	334	CCCT	393
QY	159	GGCGCGGCGACCGTGGGAGGACGACGAGGTCTCTATGAGCGTCTCTCTCTCTCTCTCTCT	218
Db	394	CGTCT	453
QY	219	CGTCT	278
Db	454	CCCT	513
QY	279	CACCGAGGCGACCGCGCGGAGCGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	338
Db	514	CCT	573
QY	339	GTTCAGGCGCCCTGCTGCGCAGACATCAAGATCGCGGCATCACCTGGCCACCGAGTTC	398
Db	574	CCCGTCT	633

## ALIGNMENTS

C	28	52.2	3.9	1280	3	US-09-060-756-4	Sequence 4, Appli
C	29	52.2	3.9	1280	4	US-09-670-314-4	Sequence 4, Appli
	30	52.2	3.9	1485	1	US-08-471-601-23	Sequence 23, Appli
	31	52.2	3.9	1485	1	US-08-474-556-23	Sequence 23, Appli
	32	52.2	3.9	1485	1	US-08-479-382-23	Sequence 23, Appli
	33	52.2	3.9	1485	1	US-08-470-354-23	Sequence 23, Appli
	34	52.2	3.9	1485	1	US-08-479-383-23	Sequence 23, Appli
	35	52.2	3.9	1485	2	US-08-479-041-23	Sequence 23, Appli
	36	52.2	3.9	1485	3	US-08-819-646-23	Sequence 23, Appli
	37	52.2	3.9	1485	4	US-09-195-716-23	Sequence 23, Appli
C	38	52.2	3.9	23673	4	US-09-773-816-1	Sequence 23, Appli
	39	52.2	3.9	30001	1	US-08-125-468-1	Sequence 1, Appli
	40	52.2	3.9	30001	2	US-08-474-933-1	Sequence 1, Appli
C	41	52.2	3.9	77536	4	US-09-410-551B-1	Sequence 1, Appli
C	42	52	3.9	614	3	US-08-998-416-861	Sequence 861, App
	43	52	3.9	1620	2	US-08-461-775-10	Sequence 10, Appli
	44	52	3.9	1620	3	US-09-031-606-10	Sequence 10, Appli
	45	51.8	3.9	4257	2	US-08-690-473-1	Sequence 1, Appli





Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103.840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 4.4%; Score 58.6; DB 3; Length 4411529;  
Best Local Similarity 44.4%; Pred. No. 0.0024;  
Matches 235; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 82 GTCCGCGGCACCGCGCTCCGCGAGCAAGCTGGGCGGCTCATCGACGGCTCCGCGCTC 141  
Db |||||  
QY 142 TTCTAAGCGCGCTCCAGCGCGGCGACCGTGGGAGCAGAGTCCCTTCTATGAGCG 201  
Db |||||  
QY 202 TTCTCGGCCACCGCGCTCCGCGCGGCGGCTGCGCGGCTGCGCGGCGGCGGCGG 3934217  
Db |||||  
QY 3934216 GGGCGCGCGCTCAGCGCGGCGGCGGCTGCGCGGCGGCGGCGGCGGCGGCGG 3934157  
Db |||||  
QY 262 GGCAGCGGCTCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3934097  
Db |||||  
QY 3934156 GTCCCGCGGCTCGCGCGGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3934037  
Db |||||  
QY 322 GACGACGCTCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3933977  
Db |||||  
QY 382 AGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3934037  
Db |||||  
QY 3934036 GCGGAGACGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3933977  
Db |||||  
QY 442 ACCAGAGCGGCTCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3933976  
Db |||||  
QY 502 GCGGTGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3933917  
Db |||||  
QY 3933916 CCCGCGGCTGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3933857  
Db |||||  
QY 562 GTGCTCGGCGGCTAGTGTATCGGTTCCGCGGCGGCGGCGGCGGCGGCGGCGG 610  
Db |||||  
QY 3933856 GCCCGGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3933808  
Db |||||

## RESULT 7

US-09-881-165-4  
; Sequence 4, Application US/09881165  
; Patent No. 6632930  
; GENERAL INFORMATION:  
; APPLICANT: HOOD, ELIZABETH  
; APPLICANT: HOWARD, JOHN  
; APPLICANT: BAILEY, MICHELE  
; APPLICANT: GASTEL, FRANS VAN  
; APPLICANT: WANG, HUAMING  
; APPLICANT: WARD, MICHAEL  
; APPLICANT: WOODARD, SUSAN

; TITLE OF INVENTION: METHOD OF INCREASING RECOVERY OF HETEROLOGOUS ACTIVE  
; FILE REFERENCE: 10032R  
; CURRENT APPLICATION NUMBER: US/09/881.165  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 60/211,732  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1082  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA encoding  
; OTHER INFORMATION: Organophosphate Hydrolase  
US-09-881-165-4

Query Match 4.3%; Score 58; DB 4; Length 1082;  
Best Local Similarity 46.2%; Pred. No. 0.0038;  
Matches 230; Conservative 0; Mismatches 265; Indels 3; Gaps 1;

QY 11 AGTCGCTACCCCTCGCGGCGGCTCCCTTCCCATACCATGTCGACCTCCACCGCG 70  
Db |||||  
QY 121 AGGCGGCTTACCTCCCTCACCGAGGACATCTCGGCTCTCCGCGGCTTCTCCGCG 180  
Db |||||  
QY 71 CGGAGCACCAAGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 130  
Db |||||  
QY 181 CTGCGCGGAGTCTTCGCGCTCCGCAAGGCGCTCGCGAGAGCGGCTCGCGGCTCC 240  
Db |||||  
QY 131 GTCGCGGCTCTTCAAGCGCTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 190  
Db |||||  
QY 241 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
Db |||||  
QY 191 TCTATGAGGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 250  
Db |||||  
QY 301 GCGAGCTGCTCTCTCGCGGAGTCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360  
Db |||||  
QY 251 CCCGCTTCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 310  
Db |||||  
QY 361 CCGGCTCTGTTGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420  
Db |||||  
QY 311 ACCTGCTCTCGAGCGCTCTCGCGGCTTCAAGGCGGCGGCGGCGGCGGCGGCGGCGG 367  
Db |||||  
QY 421 AGTTCTTCTCGGAGATCCAGTACGGCATCGAGGACACCGGCGGCGGCGGCGGCGG 480  
Db |||||  
QY 368 AGATCGGCGCATCATGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 427  
Db |||||  
QY 481 TCAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540  
Db |||||  
QY 428 CCAAGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 487  
Db |||||  
QY 541 CCCGCGCTCTCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600  
Db |||||  
QY 488 TCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 505  
Db |||||  
QY 601 ACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 618  
Db |||||

## RESULT 8

US-09-773-816-1  
; Sequence 1, Application US/09773816  
; Patent No. 6340774  
; GENERAL INFORMATION:  
; APPLICANT: Stanford University  
; APPLICANT: Khosla, Chaitan  
; TITLE OF INVENTION: NON-STEROIDAL ESTROGEN-RECEPTOR  
; FILE REFERENCE: 28600-20210.00  
; CURRENT APPLICATION NUMBER: US/09/773.816  
; CURRENT FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: US 60/243,458  
; PRIOR FILING DATE: 2000-10-25



;; PRIOR APPLICATION NUMBER: US 60/179,305  
;; PRIOR FILING DATE: 2000-01-31  
;; NUMBER OF SEQ ID NOS: 1  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 1  
;; LENGTH: 23673  
;; TYPE: DNA  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)...(23623)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-773-816-1

Query Match  
Best Local Similarity 4.3%; Score 57.6; DB 4; Length 23673;  
Matches 181; Conservative 0; Mismatches 184; Indels 3; Gaps 1;  
QY 311 ACCTCGTCTCGACGACCTCTCGCGGGTTTCAGGCGCCCTGCGTGCAGACATCAAGA 370  
DB 14307 ACCGCGCATGTGTCGGCGTGTTCGCGCTCGCGCGCTCTCGGGATCGTCCGATGC 14366  
QY 371 TCGGCGCATCACTGGCCACCGAGTTTCGCGGAGCCCTATATCGCCAAAGTGCCTCGCCA 430  
DB 14367 GCGCCACATCCCGGCCAACCGGGTCTGCTGAGACCAAGCATGATGACGTGGCGC 14426  
QY 431 AGGACCGGGACCAAGAGTTCGTCTCGGATTCGCGCTCTCGGGTCTCGAGTCTGCTCG 490  
DB 14427 CGTCTGCGCGAGCGCGCGACCTCGACCTGCTGCTCGAGTCTGCTCGCGCGCGCG 14486  
QY 491 GCGCCGAGGCGCGCTGTGCGGACGAGCGCGCGAGGTGAAGCCATGACACCGCGCG 550  
DB 14487 CCACAAAGCGCGCGCTGCGCGCTGAGCTGCGCGCGCGCGCGCGCGCGCGCGCG 14546  
QY 551 GCGTCCGCGCGTCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607  
DB 14547 GCGTCCGCGTGGGCTGTGCTGAGACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 14606  
QY 608 GTGCGTCCGCGCGGCTGTGAGAGGAAAGTGGAGTCTTGTGTCACAGTCTCGCGAGC 667  
DB 14607 ACTGCTGGCGCGCGCGTGTGAGCGCTCAAGGCGCACCGCGCTGCGCGCGCGCG 14665  
QY 668 TCAAGCGC 675  
DB 14667 TCAAGCGC 14674

RESULT 9  
US-09-105-537-32  
; Sequence 32, Application US/09105537A  
; Patent No. 6265202  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/105,537A  
; CURRENT FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 32  
; LENGTH: 11220  
; TYPE: DNA  
; ORGANISM: Streptomyces venezuelae  
US-09-105-537-32

Query Match  
Best Local Similarity 4.2%; Score 57; DB 3; Length 11220;  
Matches 222; Conservative 0; Mismatches 275; Indels 0; Gaps 0;  
QY 66 CCGCGCGAGCACCAGTTCGCGCGCGCGCTCGCGCGAGTGGGCGCGCTCAT 125

DB 8322 CCGCGCGAGCAGTTCGTTGGTGGCTGCGCGCGCGCGCGCGCGCGCGCATCTCCG 8381  
QY 126 CGACGGCTCCGGCTCTTCTACAAAGCGCTCCAGGCGCGGACCGTGGGAGACAGAGT 185  
DB 8382 CGAGGCCCTGCACGGGTGCTCGCTGCTGATGAGGCTGGTGGCGGAGCGGTTCAC 8441  
QY 186 GCGCTTATAGAGGGTTCTCCGCCACCGCGCGCGCTCCCGCGCGCGCGCGCGCAT 245  
DB 8442 CGATGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8501  
QY 246 CTTCGCCCGGTTCCAGCGCAGCGACTCTCCACCGAGCGGAGCGCGCGCGCGCGCA 305  
DB 8502 GCGGTCCACGGGACAGCGCGCGCTCTGGGCGCTCGCGCGCGCGCGCGCGCGCG 8561  
QY 306 TCCTCACCTCTCGACGACCTCTCGCGGGGTTTCAGGCGCGCTGCTGCGGAGCAT 365  
DB 8562 GGGCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8621  
QY 366 CAAGATCGGCGCCATACGTGGCCACCGAGTTCCCGGAGCGCTTACATCGCCAAAGTGC 425  
DB 8622 CGGGGACGGCTGACGACCGGGGACCGCACCGTGGCGGCGCACCTCTGAGAGCGCG 8681  
QY 426 CGCCAAAGGACCGGGGACGAGCGCTCTGCTGCGGATTCGCGCTCTCGCGCGTCCGAGT 485  
DB 8682 CGGCGCGCTCTCGCGCGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8741  
QY 486 CGTGGCGCGCGGCGCGCTGTGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8741  
DB 8742 GCGCTCTCTGTAACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8801  
QY 546 CGCGCGCGTCCGCGCGCG 562  
DB 8802 CGCGCGCGACGGCTCG 8818

RESULT 10  
US-09-105-537-5  
; Sequence 5, Application US/09105537A  
; Patent No. 6265202  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/105,537A  
; CURRENT FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 36778  
; TYPE: DNA  
; ORGANISM: Streptomyces venezuelae  
US-09-105-537-5

Query Match  
Best Local Similarity 4.2%; Score 57; DB 3; Length 36778;  
Matches 222; Conservative 0; Mismatches 275; Indels 0; Gaps 0;  
QY 66 CCGCGCGAGCACCAGTTCGCGCGCGCGCTCGCGCGAGTGGGCGCGCTCAT 125  
DB 24009 CCGCGCGACCGTCTGTTGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 24068  
QY 126 CGACGGCTCCGGCTCTTCTACAAAGCGCTCCAGGCGCGGACCGTGGGAGACAGAGT 185  
DB 24069 CGAGGCCCTGCACGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 24128  
QY 186 GCGCTTATAGAGGGTTCTCCGCCACCGCGCGCGCTCCCGCGCGCGCGCGCGCAT 245  
DB 24129 CGATGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 24188

Qy	246	CTTCCCGGGTTCA	CGGCA	CGGACTCT	CCGCA	CCGAGCGC	ACGCCCGGAGCGCA	305
Db	24189	CGCGTCCACGGG	CAGGCG	CGCTCT	GGGSCCT	CGGCGGT	TCGCGCAGACGGAGAGCC	24248
Qy	306	TCCTCACTCGT	CCTCG	ACACACT	CTCTCG	GGGGGTT	TGAGGGCCCTGGGTGCGAGACAT	365
Db	24249	GGGCGGTTCT	CTGCT	CGACT	TCGCGGGG	AGCCCG	CAGCGCGGGACGCCACGCG	24308
Qy	366	CAGATCGGCGCAT	CACGT	GGCCAC	CGAGTT	TCGCGGAGCCCT	ACATCGCCAGTGGCT	425
Db	24309	CGGGACGGGCT	TACCA	CGGGAG	CGCCAC	CCGTCGCGGCA	CCTCTGGAGACGCGCGCCT	24368
Qy	426	CGCCAAGGAC	CGCGGAC	CACAG	CGCTT	CTGTTCGGATT	TCGCGTCTCGGCGTCCGAGT	485
Db	24369	CGGACAGCGCCT	CGAC	CGCCCT	CGCT	CGGGCG	AGCCGAGCTCGCCCTCCGGAGCGG	24428
Qy	486	CGTCGCGCCCG	GGGCGCGT	TGGCGGA	CGGAG	CGCCCGGAGG	CTAAGGCCATGGACAC	545
Db	24429	GGCGTCTCT	GTATAC	CCCGGCT	TGGCGGGCC	CCGCGCGG	CGCCACGGGCTCGC	24488
Qy	546	CGCGGCGGT	TCGCGCG					
Db	24489	CGCGGCGGAC	GGGCT	CG				

```

RESULT 11
US-09-105-537-1
; Sequence 1, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikomycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 15872
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-1

```

Query Match	4.1%; Score 55.2; DB 3; Length 15872;
Best Local Similarity	44.9%; Pred. No. 0.0033;
Matches	256; Conservative 0; Mismatches 308; Indels 6; Gaps 1;
56	CCGACCTTCACCCGCCGAGACCAAGTTCGCCGGCCACCGCGCTTCGGCCAGCAGCTGG 115
3849	CCGAGCGCAACGGGCCACCCCGTGTCCGCGTGATCCGGGGCAGCGGTTCAACGGACG 3908
116	GCCCGCTCATCGAGGCTTCGGGCTCTTTTACAAGCCGCTCCAGGCCCGCGACCGTGGG 175
3909	CGCGCTCAACGGGCTCACCGCCCCAACGGGCCGCTCCCAGCAGCGCTCATCCGACGG 3968
176	AGCAGAGGTCGCGCTTATGAGGCGTTCTTCGCCCAACGCCCGCTCCCGGCCGATCC 235
3369	CCCTGGCCGACGGCGGGCTCACCCCGCCGACGTCGAGCCGTCGAGCGCACGTCACGG 4028
236	GAGACACTTCTTCCCGCGGTTCCAGGCACGGGACTCTTCCCAACGAGGCGCAGCCCG 295
4029	GTACCCGGCTCGGCGACCCCATCGAGCCGAGGCGATCTTCGGCACCTACGGCCGGGACC 4088
296	GGGAGCGCGCATCTCACTCGTCTTCGACGACCTCTCGGGGGTTTCAGGCGCCCTGCG 355
4089	GGGGCGAGGGCGCTCCGCTCCAGTCGGCTCGCTGAAATCGAACATCGGCCACGCGCAG 4148
356	TCGCGACACATCAAGATCGCGCCATCACTGTCGCCACCGAGTTTCGGCGAGCCCTATAC 415
4149	CCGCCCGGGCGTGGCGGGCTCTACAGATGGTCTTCGCGATCGGCCACGGCGTCTCTGC 4208

[illegible]

```

RESULT 12
US-09-091-609-1
; Sequence 1, Application US/09091609
; Patent No. 6600029
; GENERAL INFORMATION:
; APPLICANT: SHERMAN, DAVID H.
; APPLICANT: WILLIAMS, MARK D.
; APPLICANT: XUE, YONGQUAN
; TITLE OF INVENTION: METABOLIC ENGINEERING OF
; TITLE OF INVENTION: POLYHYDROXYALKANOATE MONOMER SYNTHASES
; FILE REFERENCE: 600.297US2
; CURRENT APPLICATION NUMBER: US/09/091.609
; CURRENT FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: PCT/US96/20119
; EARLIER FILING DATE: 1996-12-18
; EARLIER APPLICATION NUMBER: 60/008,847
; EARLIER FILING DATE: 1995-12-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 15872
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)...(13909)
US-09-091-609-1

```

Query Match	4.1%;	Score 55.2;	DB 4;	Length 15872;
Best Local Similarity	44.9%;	Pred. No. 0.0033;		
Matches 256;	Conservative 0;	Mismatches 308;	Indels 6;	Gaps 1;
56	CGGACCTTCCACCGCGGAGACCAAGTCCGCGGCACCGGCGCTCCGACGACGAGCTGG	115		
3849	CCGAGCGACACGCGCCACCCCGTGCTCGCGGTATCGGGGCGACGCGGTCAACGAGGACG	3908		
116	GCCGCGCTCATCGAGGCGCTCGGCGCTTTTACAAGCGGCTCCAGGCCGCGGACCGTGGG	175		
3909	GCGCTCTCCAAAGGGGTCAACGCGCCCAAGGCCCTCCAGAGCGCGTCTATCCGACAGG	3968		
176	AGCACAGAGTGGCTTCTAAGAGGGTCTTCGCGCCACCGCGCGCGTCCCGGCCGCGATCC	235		
3969	CCCTGGCGACGCGGGGCTCACCCGCGCGAGTCTGACGCCGTGAGGCGGCACGTACGG	4028		
236	GAGACACTTCTTCCCGCGTTCACGGGACGCGACTCTCTCCCAACGAGGCGGACGCCG	295		
4029	GTACCGGCTCGCGCACCCCATCGAGGCGGAGGGGATCTCTGGGACCTACGGCGGGACC	4088		
296	GGGAGCGCGATCCTCACCTGTGCTCTGACGACTCTCTCGGGGGTTTACGCGCCCTGCG	355		
4089	GGGGCGAGGCGCTCCGCTTCCAGCTCGGCTCGCTGAAGTGGAAATCGGCGCACGCGCAGG	4148		
356	TGCGAGACATCAGATCCGGCGGCATACGTGGCCACCGAGTTTCGCGGAGCGCTTACATCG	415		
4149	CCGCGCGGGCGTGGCGGGGCTCATCAGATGTGTCCTCGGATCGCGCACGGGGTCTCTGC	4208		

QY 416 CCAAGTGCCTCGCCAGGACCGCGGACCAAGAGCGTTCTGCTCGGATTCCGGTCTCCG 475  
Db |||||  
QY 4209 CCAGGAGCTCCACGTGAGCGGCCACCAACCCCGC-----TCGACTGGAGGCGCGCG 4262  
Db |||||  
QY 476 GCCTCCGAGTCTCGGCGCCCGAGGCGCGCTGTGGCGGACGAGCGCGCCGAGTGAAG 535  
Db |||||  
QY 4263 GCCTCGAGTCTCTACCGAGGAGCGAGTGGCGCGGAGACGGCGCGCGCGCGCGCG 4322  
Db |||||  
QY 536 CCATGGACACCGCGCGCTCGCGCGCTGTCTCGGCGCTACGTGTCTATCTCGTTCGCGACG 595  
Db |||||  
QY 4323 CGATCTCTCTCTCGGCGATCAGCGGACCAAGCGCCACATCGTGGTGGACAGCGCCCGCG 4382  
Db |||||  
QY 596 AGGGATGAGTGTGCGCTCGCGCGCGCG 625  
Db |||||  
QY 4383 AGCGGGGAGCGCGGTCAACACACCG 4412  
Db |||||

## RESULT 13

US-09-091-609-3  
; Sequence 3, Application US/09091609  
; Patent No. 6600029  
; GENERAL INFORMATION:  
; APPLICANT: SHERMAN, DAVID H.  
; APPLICANT: WILLIAMS, MARK D.  
; APPLICANT: XUE, YONGQUAN  
; TITLE OF INVENTION: METABOLIC ENGINEERING OF  
; FILE REFERENCE: POLYHYDROXYALKANOATE MONOMER SYNTHASES  
; CURRENT APPLICATION NUMBER: US/09/091,609  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER FILING DATE: PCT/US96/20119  
; EARLIER FILING DATE: 1996-12-18  
; EARLIER FILING DATE: 1995-12-19  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 15872  
; TYPE: DNA  
; ORGANISM: Streptomyces venezuelae  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (14148)....(15824)  
US-09-091-609-3

Query Match 4.1%; Score 55.2; DB 4; Length 15872;  
Best Local Similarity 44.9%; Pred. No. 0.0033;  
Matches 256; Conservative 0; Mismatches 308; Indels 6; Gaps 1;  
QY 56 CCGACTCCACCGCGCGGAGCACCAAGTCCGCGGCGCACCGCGCTTCGCGCAAGCTGG 115  
Db |||||  
QY 3849 CCGAGCGCAACGCCACCCCGTCTCGCGTGTATCCGGGCGAGCGCGGTCAACCGAGCG 3908  
Db |||||  
QY 116 GCGCGTCAATACAGCGCTCCGCGCTTCTCAAGCGCGTCCAGGCGCGGCGCGCGTGGG 175  
Db |||||  
QY 3909 GCGCTCCAGCGGTCAACCGCGCCCAACGCGCGCTCCAGAGCGCGTCAATCCGACAG 3968  
Db |||||  
QY 176 AGCAGAGTCCCTTCTATGAGGCGTCTCTCGCGCCACGCGCGCTCCGCGCGCGCATCC 235  
Db |||||  
QY 3969 CCTGCGGACGCGCGGCTACCGCGCGCGAGTGAAGCGCTTGAAGGCGGACGATACG 4028  
Db |||||  
QY 236 GAGACACTTCTTCCCGGTTCCAGGCAACGCGACTCTCTCCCGACCGAGGCGCGCGCG 295  
Db |||||  
QY 4029 GTACCGGCTCGCGACCCCATCGAGGCGAGGCGATCTCTCGGCACTACGCGCGGACC 4088  
Db |||||  
QY 296 GGGAGCGCATCTCACTCGTCTTCGAGCACTCTCTCGCGGGTTTCAGGCGCGCTTCG 355  
Db |||||  
QY 4089 GGGGAGGCGGCTCCGCTTCAAGCTCGGCTCGGCTGAAATCGAATCGGCCACGCGGAG 4148  
Db |||||  
QY 356 TCGCAGACATCAAGATCGCGCGCATCACGTGGCCACCGAGTTTCGCGGAGCGCTTACATCG 415  
Db |||||  
QY 4149 CCGCGCGGCGTGGCGGCTCATCAGATGGTCTCTCGGATGCGCCACGCGCTCTCG 4208  
Db |||||

QY 416 CCAAGTGCCTCGCCAGGACCGCGGACCAAGAGCGTTCTGCTCGGATTCCGGTCTCCG 475  
Db |||||  
QY 4209 CCAGGAGCTCCACGTGAGCGGCCACCAACCCCGC-----TCGACTGGAGGCGCGCG 4262  
Db |||||  
QY 476 GCCTCCGAGTCTCGGCGCCCGAGGCGCGCTGTGGCGGACGAGCGCGCCGAGTGAAG 535  
Db |||||  
QY 4263 GCCTCGAGTCTCTACCGAGGAGCGAGTGGCGCGGAGACGGCGCGCGCGCGCGCG 4322  
Db |||||  
QY 536 CCATGGACACCGCGCGCTCGCGCGCTGTCTCGGCGCTACGTGTCTATCTCGTTCGCGACG 595  
Db |||||  
QY 4323 CGATCTCTCTCTCGGCGATCAGCGGACCAAGCGCCACATCGTGGTGGACAGCGCCCGCG 4382  
Db |||||  
QY 596 AGGGATGAGTGTGCGCTCGCGCGCGCG 625  
Db |||||  
QY 4383 AGCGGGGAGCGCGGTCAACACACCG 4412  
Db |||||

## RESULT 14

US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: TUBERCULOSIS  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 4.1%; Score 55.2; DB 3; Length 4403765;  
Best Local Similarity 43.4%; Pred. No. 0.014; Mismatches 328; Indels 0; Gaps 0;  
Matches 252; Conservative 0; Mismatches 328; Indels 0; Gaps 0;  
QY 46 CATACCATGTCCGACTCCACCGCGCGGAGCACCAAGTCCGCGGCGCACCGCGCTCCGCG 105  
Db |||||  
QY 150482 CAGCCCAACCGCGCGCACCGCGCGCGCATGCTGCTGACCGGCGAGCGGTGGGCGCC 150423  
Db |||||  
QY 106 AGCAGCTGGGCGCGCTCATCGAGCTCCGCGCTTCTTACAGCGGTCCAGGCGCGC 165  
Db |||||  
QY 150422 GTACCAACCGCTTCCCGCGGTGTCCCGCGCGGTGTCCCGCGCGCGAGTACCCCAAGTCCGCG 150363  
Db |||||  
QY 166 GACCTGGGAGCACGAGTCTGCTTCTATGAGGCGTTCGCGCGCACGCGCGCTCCGCG 225  
Db |||||  
QY 150362 GCGCGCGCTTGGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGTCC 150303  
Db |||||  
QY 226 GCGCGCATCCGAGACACTTCTTCCCGCGGTTCACAGCGGACGCGACTCTCTCCCAACCGAG 285  
Db |||||  
QY 150302 GCGGTGTCCCGCGCGCGCGCTGCGCGTAGAGCGCATCCGCGCACCGCGCACCGCGC 150243  
Db |||||  
QY 286 GCGCAGCCCGGAGCGCGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 345  
Db |||||  
QY 150242 GAGGCACCGTCTCCACCGCTGACGATGTTGGCATCGGCGCGCTTCCGCGCATTCGCCACC 150183  
Db |||||  
QY 346 GCGCGCTCTGCTCGCAGACATCAAGATCGGCGCGCATCACTGCGCGCACCGAGTTTCGCGGAG 405  
Db |||||  
QY 150182 GCG 150123  
Db |||||  
QY 406 CCTACATCGCCAAAGTGCCTCGCCAAAGACCGCGGAGCACGAGCGTTCGTCTCGGATTC 465  
Db |||||

Db 150122 GCCTGGCCGACACCGCCCGCGGATCGCCGATCGCCGCGCTGTCGCCACAGCCATCC 150063  
QY 466 CGGCTTCGCGGTCGCGAGTCGTCGCGCCCGAGCGCGCGCTGTGGCGAGCGGCGCCG 525  
Db 150062 GCGGTGCG 150003  
QY 526 GAGGTGAAGCCATGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 585  
Db 150002 GCGGATCAGCCCG 149943  
QY 586 GTTCCGACGAGGGGATGACTGCTGCGCTCGCGCGCGCGCG 625  
Db 149942 ATTGCGCCATTGCCGTACAGATCCCCCAGCGCCGCGCG 149903

## RESULT 15

US-08-528-199-2  
; Sequence 2, Application US/08528199  
; Patent No. 5763228  
; GENERAL INFORMATION:  
; APPLICANT: KUBOTA, Michio  
; APPLICANT: TSUSAKI, Keiji  
; APPLICANT: SUGIMOTO, Toshiyuki  
; TITLE OF INVENTION: RECOMBINANT ENZYME FOR CONVERTING  
; TITLE OF INVENTION: MALTOSE INTO TREHALOSE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/528,199  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/485,126  
; FILING DATE: 14-SEP-1995  
; APPLICATION NUMBER: US 08/485,126  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 156399/1994  
; FILING DATE: 16-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROWDY, Roger L.  
; REGISTRATION NUMBER: 25,618  
; REFERENCE/DOCKET NUMBER: KUBOTA-6A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; TELEX: 248633  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1704 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-528-199-2

Query Match 4.1%; Score 54.8; DB 1; Length 1704;  
Best Local Similarity 44.3%; Pred. No. 0.0022;  
Matches 224; Conservative 0; Mismatches 282; Indels 0; Gaps 0;  
QY 73 GAGCACCAGTCGCGCGCCACCGCGCTCCGCGCAGCAGCTGGCGCGCTCATCGACGCG 132

Db 199 GGTAGCAGCTGCGCGACTACACCGGATCTCTCCGAGATCGGCACTGGTCGAGACTTC 258  
QY 133 TCCGGCTCTTCTTACAAGCCGCTCCAGGCGCGCGACCGTGGGAGACACGAGGTGCGCTTC 192  
Db 259 CAGCGCTTCTCTGACGGCGCGACGAGCGCGGGATCCGGGTGATCATCGACTTCGTCATG 318  
QY 193 TATGAGGCGTTCCTCCGCCACACCCCGCGCTCCCGGCCCGCATCCGAGACACCTTCTTCCCG 252  
Db 319 AACCAACAGAGTACGCGCACCGCTGTTCAGGCTCTCCGCGAGCATCCCGACGCGCG 378  
QY 253 CGGTTCCACGCGCAGCGACTCTCTCCCGACCGAGGCGCGCGCGGAGCGCGATCTCTAC 312  
Db 379 TACGCGGACTTCTACGTCTGTCTCCGACACCGACGAGCTTACCGAGACGCGGGGTGATC 438  
QY 313 CTCGTCTCTGACGACCTCTCTCGCGGGTTTCAGGCGCGCTGCTCGCAGACATCAAGATC 372  
Db 439 TTGCTCGACACCGAGCCGCTCGAATCTGAGCGTGGGACAGACCCGCGGCGCTACTCTGG 498  
QY 373 GCGCCATCAGTGGCCACCGAGTTCGCGGAGCCCTTACATCGGCAAGTGCCTCGCCAAAG 432  
Db 499 CACGCTTCTTCCACACCGAGCCGACCTGAATTCGACAAACCGGAGGTCCAGGACGCC 558  
QY 433 GACCGCGGACACGAGCGTTCTCTCGGATTTCGCGCTCTCCGGGCTCCGAGTGTCTGCGC 492  
Db 559 ATGCTGGAGGCGATGGCGTCTGTGCTGACATGGGCTTCGACGCTTCCGGCTCGACGCG 618  
QY 493 CCCAGGGCGCGTGTGCGGACGAGCGCGCGCGGAGGTGAAGGCCATGACACCGCGCGC 552  
Db 619 GTGCCCTACCTCTACGAGCGCTCCCGCACCAAGCGCGAGAACCTCCCCGAGACGACGAG 678  
QY 553 GTCCGCGCGCTGCTCCCGCGCTAGCT 578  
Db 679 ATGCTCAAGCGGGTGGCGCGCTTCTGT 704

Search completed: March 27, 2004, 04:39:20

Job time : 150 secs

GenCore version 5.1.1.6  
 Copyright (c) 1993 - 2004 CompuGen Ltd.  
 OM nucleic - nucleic search, using sw model  
 Run on: March 27, 2004, 03:33:25 ; Search time 513 Seconds  
 (without alignments)  
 9753.850 Million cell updates/sec

Title: US-10-042-894A-7  
 Perfect score: 1344  
 Sequence: 1 GCACGAGTGCCTCGTCAC.....ataaaaaaaaaaaaaa 1344

Scoring table:

IDENTITY NUC  
 Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4517892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
17:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
18:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	1344	14	US-10-042-894A-7
2	865.4	64.4	923	14	US-10-042-894A-5
3	844.8	62.9	1169	14	US-10-042-894A-1
4	837.4	62.3	923	14	US-10-042-894A-3
5	740.4	55.1	3416	14	US-10-042-894A-20
6	686	51.0	899	14	US-10-042-894A-15
7	470.4	35.0	643	14	US-10-042-894A-17
8	298.4	22.2	519	14	US-10-042-894A-18
9	191.6	14.3	353	14	US-10-042-894A-19
10	184.6	13.7	1195	14	US-10-042-894A-11
11	179	13.3	1105	14	US-10-042-894A-9
12	179	13.3	1345	12	US-10-424-599-106467
13	149.8	11.1	1020	14	US-10-042-894A-13
14	104.6	7.8	464	9	US-09-770-444-340
15	71.4	5.3	135638	14	US-10-314-657-1

16	68.6	5.1	1383	12	US-10-417-700A-10
17	68.6	5.1	88421	9	US-09-976-059-1
18	68.2	5.1	2541	14	US-10-156-761-2855
19	68.2	5.1	9025608	14	US-10-156-761-1
20	66.8	5.0	1347	12	US-10-424-599-118038
21	66.2	4.9	3471	14	US-10-156-761-3726
22	65.4	4.9	484	10	US-09-770-961-829
23	64.8	4.8	825	14	US-10-156-761-4706
24	64.8	4.8	9025608	14	US-10-156-761-1
25	64.2	4.8	1296	12	US-10-425-114-33515
26	63.8	4.7	5452	14	US-10-017-161-1481
27	63.8	4.7	5452	15	US-10-292-798-1189
28	63.6	4.7	987	14	US-10-156-761-3940
29	63.4	4.7	1581	14	US-10-156-761-3431
30	61.8	4.6	7788	14	US-10-329-079-8
31	61.8	4.6	11058	14	US-10-156-761-3629
32	61.8	4.6	37360	14	US-10-329-079-6
33	61.4	4.6	18435	14	US-10-156-761-412
34	61.4	4.6	100000	14	US-10-156-761-15103
35	61.2	4.6	2475	15	US-10-260-238-5598
36	61.2	4.5	2662	12	US-10-425-114-16619
37	61	4.5	1271	12	US-10-425-114-1413
38	61	4.5	1385	12	US-10-425-114-4294
39	61	4.5	2465	15	US-10-369-493-42318
40	61	4.5	64492	12	US-10-378-083-1
41	60.2	4.5	1266	14	US-10-145-413-104
42	60.2	4.5	2291	12	US-10-425-114-16568
43	60	4.5	1465	12	US-10-425-114-27827
44	60	4.5	1500	12	US-10-425-114-28844
45	59.8	4.4	1440	14	US-10-156-761-6190

## ALIGNMENTS

## RESULT 1

US-10-042-894A-7  
 ; Sequence 7, Application US/10042894A  
 ; Publication No. US20030009011A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shi, Jinrui  
 ; APPLICANT: Beach, Larry  
 ; APPLICANT: Wang, Hongyu  
 ; APPLICANT: Rafalski, Antoni J.  
 ; APPLICANT: Cahoon, Rebecca E.  
 ; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase  
 ; FILE REFERENCE: 1286  
 ; CURRENT APPLICATION NUMBER: US/10/042,894A  
 ; CURRENT FILING DATE: 2002-01-09  
 ; PRIOR APPLICATION NUMBER: US 60/261,465  
 ; PRIOR FILING DATE: 2001-01-12  
 ; NUMBER OF SEQ ID NOS: 37  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 1344  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (52)...(921)  
 ; US-10-042-894A-7

Query Match 100.0%; Score 1344; DB 14; Length 1344;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCACGAGTGCCTCGTCACCCCTCGGCCCATAGTCCCTTCCCATACCATGTCGCAC 60  
 Db 1 GCACGAGTGCCTCGTCACCCCTCGGCCCATAGTCCCTTCCCATACCATGTCGCAC 60  
 QY 61 CTCACCCGCGGAGCACCAAGTCGCGGCCACCGCGCCCTCCGCGAGCTGGGCGCG 120

```
Db 61 CTCACCCCGCGAGACACCAAGTCCGCGGCACCGCGCTCCGCCAGCAAGCTGGGCCG 120
Qy 121 CTCATCGAGGGTCCGGGCTCTTCTACAGCCGCTCCAGGCGGCGCAACGTTGGGAGAC 180
Db 121 CTCATCGAGGGTCCGGGCTCTTCTACAGCCGCTCCAGGCGGCGCAACGTTGGGAGAC 180
Qy 181 GAGTTCGGCTTCTATGAGCGTTCTCCGCCACCGCGCTCCGGCGCGGATCCGAGAC 240
Db 181 GAGTTCGGCTTCTATGAGCGTTCTCCGCCACCGCGCTCCGGCGCGGATCCGAGAC 240
Qy 241 ACCTTCTTCCCGCGGTTCCAGGCAAGCGATCTCTCCCAACGAGGCGAGCCGCGGAG 300
Db 241 ACCTTCTTCCCGCGGTTCCAGGCAAGCGATCTCTCCCAACGAGGCGAGCCGCGGAG 300
Qy 301 CGCATCTCACTCGTCTCGAGACCTCTCGGGGTTTCAGGCGCTTGGTTCGCA 360
Db 301 CGCATCTCACTCGTCTCGAGACCTCTCGGGGTTTCAGGCGCTTGGTTCGCA 360
Qy 361 GACATCAAGATCGGCGCATCACGTGGCCACCGAGTTCGCGGAGCCCTACATGCCAAG 420
Db 361 GACATCAAGATCGGCGCATCACGTGGCCACCGAGTTCGCGGAGCCCTACATGCCAAG 420
Qy 421 TGCCTCGCAAGACCGCGGACACAGAGGTTCTCTCGATTCGCGGTC 480
Db 421 TGCCTCGCAAGACCGCGGACACAGAGGTTCTCTCGATTCGCGGTC 480
Qy 481 CGAGTCGTCGCGCGGCGCGTGTGGCGGACGAGCGCCCGAGGTTAAAGCCATG 540
Db 481 CGAGTCGTCGCGCGGCGCGTGTGGCGGACGAGCGCCCGAGGTTAAAGCCATG 540
Qy 541 GACACCGCGCGCTCCGCGCGTGTCTCGGCGCTACGTGTATCCGTTGCCAGAGGG 600
Db 541 GACACCGCGCGCTCCGCGCGTGTCTCGGCGCTACGTGTATCCGTTGCCAGAGGG 600
Qy 601 ATGACTGTGCGCTCGCGCGGCGGTGTAACGAGAAAGGTGAGTCTTGTACAGCTG 660
Db 601 ATGACTGTGCGCTCGCGCGGCGGTGTAACGAGAAAGGTGAGTCTTGTACAGCTG 660
Qy 661 CGCAGCTCAAGGCGTGTTCAGAGACAGACTCTGTTCACATCTTACTGGGTCGATT 720
Db 661 CGCAGCTCAAGGCGTGTTCAGAGACAGACTCTGTTCACATCTTACTGGGTCGATT 720
Qy 721 CTTCTGGGCTATGATCTGTCTGACGACGAGGAGATGGGGTGGGGTGACGGTG 780
Db 721 CTTCTGGGCTATGATCTGTCTGACGACGAGGAGATGGGGTGGGGTGACGGTG 780
Qy 781 AAGCTGTGACTTTCGCGCATGTGGCGAGGAGTGTGGGGTGAATGACCAACTTCCTG 840
Db 781 AAGCTGTGACTTTCGCGCATGTGGCGAGGAGTGTGGGGTGAATGACCAACTTCCTG 840
Qy 841 GGGGGCTCTGCTCGCTGATCAAGTTCGTTTCTGACATTTTCGAGAGACTCTCTATAC 900
Db 841 GGGGGCTCTGCTCGCTGATCAAGTTCGTTTCTGACATTTTCGAGAGACTCTCTATAC 900
Qy 901 CAGCTTTGGGCTCTTCTTAAGAGAGGATCTGSCATTTTCGATTTGATAACAAAGCCCTA 960
Db 901 CAGCTTTGGGCTCTTCTTAAGAGAGGATCTGSCATTTTCGATTTGATAACAAAGCCCTA 960
Qy 961 CAGTTTGTCTGAAAGAGAGGCGCTCCGAGTTGTGCTGGGTGTGAGATCTGAGAGC 1020
Db 961 CAGTTTGTCTGAAAGAGAGGCGCTCCGAGTTGTGCTGGGTGTGAGATCTGAGAGC 1020
Qy 1021 GTGTCGCGCCACTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1080
Db 1021 GTGTCGCGCCACTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1080
Qy 1081 CTTTTCGCAACCCCTTACTTCGAGAGAACTTTTTCCTTCCCACTTTGGGGTTCGA 1140
Db 1081 CTTTTCGCAACCCCTTACTTCGAGAGAACTTTTTCCTTCCCACTTTGGGGTTCGA 1140
Qy 1141 TTACGTTGATCTGTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 1200
Db 1141 TTACGTTGATCTGTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 1200
```

## RESULT 2

```
US-10-042-894A-5
; Sequence 5, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalaki, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 923
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(922)
US-10-042-894A-5
```

```
Query Match 64.4%; Score 865.4; DB 14; Length 923;
Best Local Similarity 98.8%; Pred. No. 1.1e-222;
Matches 870; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```
Qy 42 TCCCCATACCATGTCCGACCTCCACCGCGCGGAGCACCAAGTCGCGGCGCACCGCGCTC 101
Db 43 TCCCCATACCATGTCCGACCTCCACCGCGCGGAGCACCAAGTCGCGGCTCACCGGCTC 102
Qy 102 CGCCAGCAAGCTGGGCCCGCTCATCGAGCGCTCCGGCTCTTCTCAAGCGCTCCAGGC 161
Db 103 CGCCAGCAAGCGGGCGCGCTCATCGAGCGCTCCGGCTCTTCTCAAGCGCTCCAGGC 162
Qy 162 CGCGGACCGTGGGAGCAGAGTCCCTTCTATGAGGCTTCTCGGCCACCGCGCT 221
Db 163 CGCGGACCGTGGGAGCAGAGTCCCTTCTATGAGGCTTCTCGGCCACCGCGCT 222
Qy 222 CCGGGCCCGCATCCGAGACACCTTCTTCCCGGTTTCCACGCGCAGCGACTCTCCGCCAC 281
Db 223 CCGGGCCCGCATCCGAGACACCTTCTTCCCGGTTTCCACGCGCAGCGACTCTCTCCGCCAC 282
Qy 282 CGAGGCGAGCGCGGAGCGCGCATCTCTCACTCTGCTCTCGAGCGCTCTCTCGCGGGTT 341
Db 283 CGAGGCGAGCGCGGAGCGCGCATCTCTCACTCTGCTCTCGAGCGCTCTCTCGCGGGATT 342
Qy 342 TCAGGCGCGCTCGTCCGAGATCAAGATCGCGGCGCATCACTGCGGCGAGGTTGCGC 401
Db 343 TCAGGCGCGCTCGTCCGAGATCAAGATCGCGGCGCATCACTGCGGCGAGGTTGCGC 402
Qy 402 GGAGCCCTACATCGCCAGTGTCTGCCAAGACCGCGGAGCACAGAGGTTCTGCTCG 461
Db 403 GGAGCCCTACATCGCCAGTGTCTGCCAAGTGTCTGCCAAGACCGCGGAGCACAGAGGTTCTGCTCG 462
```





;; PRIOR APPLICATION NUMBER: US 60/261,465  
;; PRIOR FILING DATE: 2001-01-12  
;; NUMBER OF SEQ ID NOS: 37  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 3  
;; LENGTH: 923  
;; TYPE: DNA  
;; ORGANISM: Zea mays  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (53)...(736)  
US-10-042-894A-3

Query Match 62.3%; Score 837.4; DB 14; Length 923;  
Best Local Similarity 97.5%; Pred. No. 3.8e-215;  
Matches 859; Conservative 0; Mismatches 21; Indels 1; Gaps 1;  
QY 42 TCCCATACCATGTCGACCTCCACCCGCGGAGCACCAAGTCGCGGCCACCGCGCTC 101  
DB 43 TCCCATACCATGTCGACCTCCACCCGCGGAGCACCAAGTCGCGGCCACCGCGCTC 102  
QY 102 CGCAGCAAGTGGGCGGCTCATCGAGGCTCGGCTCTTCTAAGCGGCTCCAGGC 161  
DB 103 CGCAGCAAGTGGGCGGCTCATCGAGGCTCGGCTCTTCTAAGCGGCTCCAGGC 162  
QY 162 CGCGACCGTGGGAGCACGAGTGGCTTCTATGAGCGTTCCTCCGCCACCGCGCGT 221  
DB 163 CGCGACCGTGGGAGCACGAGTGGCTTCTATGAGCGTTCCTCCGCCACCGCGCGT 222  
QY 222 CCAGCGCGCATCCGAGACACTTCTTCCCGGTTCCAGGACGCGACTCTCCCGAC 281  
DB 223 CCAGCGCGCATCCGAGACACTTCTTCCCGGTTCCAGGACGCGACTCTCCCGAC 282  
QY 282 CGAGCGGAGCCCGGAGCGCGATCTCACTCTCTCGACGACTCTTCGCGGATT 341  
DB 283 CGAGCGGAGCCCGGAGCGCGATCTCACTCTCTCGACGACTCTTCGCGGATT 342  
QY 342 TCAGCGCGCTCGCTCGCAGACATCAAGATCGGCGCATCACTGCGGCGCGCGAGTTCGCC 401  
DB 343 TGAGCGCGCTCGCTCGCAGACATCAAGATCGGCGCATCACTGCGGCGCGCGAGTTCGCC 402  
QY 402 GGAGCGCTTACATCGCCAGTGCCTCGCCAGGACCGCGGACCAAGCGGCTTCGTCGG 461  
DB 403 GGAGCGCTTACATCGCCAGTGCCTCGCCAGGACCGCGGACCAAGCGGCTTCGTCGG 462  
QY 462 ATTCCGGCTCTCGCGCTCGAGTGGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 521  
DB 463 ATTCCGGCTCTCGCGCTCGAGTGGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 522  
QY 522 CCGGAGGTGAAGGCATGGACACCGCGCGCGTCCGCGCGTCTCGCGCGCTACGTGTC 581  
DB 523 CCGGAGGTGAAGGCATGGACACCGTGGGCTCGCGCGTCTCGCGCGCTACGTGTC 582  
QY 582 ATCCG-TTGCAGAGGGATGAGTGTGGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 640  
DB 583 ATCCG-TTGCAGAGGGATGAGTGTGGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 642  
QY 641 GTGGAGTCTGTCAAGCTCGCGAGCTCAAGGCTGGTTCGAGGAGCAGACTCTGTTC 700  
DB 643 GTGGAGTCTGTCAAGCTCGCGAGCTCAAGGCTGGTTCGAGGAGCAGACTCTGTTC 702  
QY 701 ACTTCTACTCGCGCTCGATTTCTTGGGCTATGATGCTGCTGCTCGAGTCGAGCGGAG 760  
DB 703 ACTTCTACTCGCGCTCGATTTCTTGGGCTATGATGCTGCTGCTCGAGTCGAGCGGAG 762  
QY 761 ATGGGGTGGGTGACGAGTGGTGGATTTTGGCCATGTCGCCAGGGTGTATGGG 820  
DB 763 GTGGGGTGGGTGACGAGTGGTGGATTTTGGCCATGTCGCCAGGGTGTATGGG 822  
QY 821 TGATTTACCAACAATCTCTGGCGGCTCTGCTCGCTGATCAAGTTCCTTCTGACATTG 880  
DB 823 TGATTTACCAACAATCTCTGGCGGCTCTGCTGCTGATCAAGTTCCTTCTGACATTG 882

QY 881 TTCCGAGACTCTCTCATAGCAGCCCTTGGGTCTTCTTAA 921  
DB 883 TTCCGAGACTCTCTCATAGCAGCCCTTGGGTCTTCTTAA 923

## RESULT 5

US-10-042-894A-20  
;; Sequence 20, Application US/10042894A  
;; Publication No. US20030009011A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Shi, Jinrui  
;; APPLICANT: Beach, Larry  
;; APPLICANT: Wang, Hongyu  
;; APPLICANT: Rafalski, Antoni J.  
;; APPLICANT: Cahoon, Rebecca E.  
;; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase  
;; FILE REFERENCE: 1286  
;; CURRENT APPLICATION NUMBER: US/10/042,894A  
;; CURRENT FILING DATE: 2002-01-09  
;; PRIOR APPLICATION NUMBER: US 60/261,465  
;; PRIOR FILING DATE: 2001-01-12  
;; NUMBER OF SEQ ID NOS: 37  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 20  
;; LENGTH: 3416  
;; TYPE: DNA  
;; ORGANISM: Zea mays  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (72)...(407)  
US-10-042-894A-20

Query Match 55.1%; Score 740.4; DB 14; Length 3416;  
Best Local Similarity 90.2%; Pred. No. 6.4e-189;  
Matches 862; Conservative 0; Mismatches 31; Indels 63; Gaps 4;

QY 2 CACGAGTCAGTCGTCACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCGCGACC 61  
DB 28 CACCATCGCCACCGTCCACCCCTTGTCCCATAG-----TCCCATACCATGTCGCGACC 81  
QY 62 TCACCCCGCGGAGCACCAAGTCGCGGCCACCGCGCTCGGCGAGCAAGCTGGCGCGC 121  
DB 82 TCACCCCGCGGAGCACCAAGTCGCGGCCACCGCGCTCGGCGAGCAAGCTGGCGCGC 141  
QY 122 TCATCGAGCGCTCGGCGCTCTTCAAGCGCTTCAAGCGCGCTCCAGGCGCGGACCGTGGGAGCACG 181  
DB 142 TCATCGAGCGCTCGGCGCTCTTCAAGCGCGCTCCAGGCGCGGACCGTGGGAGCACG 201  
QY 182 AGGTGCGCTTCTATGAGCGCTTCTCGGCCACCGCGCGCTCGGCGCGCTCCGAGACA 241  
DB 202 AGGTGCGCTTCTATGAGCGCTTCTCGGCCACCGCGCGCTCCGCGCGCTCCGAGACA 261  
QY 242 CTTCTTTCCTCGGCTTCAGGCGACCGAGCTCTTCCCGCGCGCGCGCGCGCGCGCG 301  
DB 262 CTTCTTTCCTCGGCTTCAGGCGACCGAGCTCTTCCCGCGCGCGCGCGCGCGCGCG 321  
QY 302 CGCATCTCTCACTCGTCTCTCGAGCGACTCTCGCGGCTTTCAGGCGCGCTTCGCGAG 361  
DB 322 CGCATCTCTCACTCGTCTCTCGAGCGACTCTTCCCGCGCGCGCGCGCGCGCGCG 381  
QY 362 ACATCAAGATCGGCGCGCATCACTGCGGCGACCGAGTTCGCGGAGCCCTTACATGCCAAGT 421  
DB 382 ACATCAAGATCGGCGCGCATCACTGCGGCGACCGAGTTCGCGGAGCCCTTACATGCCAAGT 405  
QY 422 GCCTCGCAAGGACCGCGGAGCACGAGCGTTCGCTCGGATTCGCGCTTCGCGCGCTCC 481  
DB 406 -----GACCACGAGCGTTCGCTCGGATTCGCGCTTCGCGCGCTCC 446  
QY 482 GAGTCTCGGCGCGGCGCGCTGTGGCGAGCGGCGCGCGCGCGCGCGCGCGCGCGCG 541  
DB 447 GAGTCTCGGCGCGGCGCGCTGTGGCGAGCGGCGCGCGCGCGCGCGCGCGCGCGCG 506



```

; LENGTH: 643
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(643)
; OTHER INFORMATION: n = A, T, C or G
US-10-042-894A-17

Query Match      35.0%; Score 470.4; DB 14; Length 643;
Best Local Similarity 95.9%; Pred. No. 1.8e-116;
Matches 509; Conservative 0; Mismatches 18; Indels 4; Gaps 3;

QY 42 TCCCCATACATGTCGAGCTCCACCGCGCGGAGACCAAGTCCGCGGCCACCGCGCCCTC 101
Db 107 TCCCCATACATGTCGAGCTCCACCGCGCGGAGACCAAGTCCGCGGCCACCGCGCCCTC 166
QY 102 CGCCAGCAAGCTGGCGCGCTCATGACGGCTCCGGCTCTTCTACAAGCCGCTCCAGGC 161
Db 167 CGCCAGCAAGCTGGCGCGCTCATGACGGCTCCGGCTCTTCTACAAGCCGCTCCAGGC 226
QY 162 CGCGACCGCTGGGAGCAGAGGTGCGCTTCTATGAGGGCTTCTCCGCCACCGCGCGGT 221
Db 227 CGCGACCGCTGGGAGCAGAGGTGCGCTTCTATGAGGGCTTCTCCGCCACCGCGCGGT 286
QY 222 CGCGCGCGCTATCGAGACACCTTCTCCCGCGGTCCAGGCGACCGGACTCTCCCGAC 281
Db 287 CGCGCGCGCTATCGAGACACCTTCTCCCGCGGTCCAGGCGACCGGACTCTCCCGAC 346
QY 282 CGAGGCGCAGCCCGGAGCGCATCTCACCTCTGCTCGACGACCTCTCCGCGGGGT 341
Db 347 CGAGGCGCAGCCCGGAGCGCATCTCGACCTCTGCTCGACGACCTCTCCGCGGGGT 406
QY 342 TCAGGCGCCCTGCGTCCGACATCAAGATCGCGCGCATCACGTTGCGCACCGAGTTGCGC 401
Db 407 TGAGGCGCCCTGCGTCCGACATCAAGATCGCGCGCATCACGTTGCGCACCGAGTTGCGC 466
QY 402 GAGGCGCTATCGCCAAAGTGCCTCGCCAAAGACCGCGGACCAAGAGCGTTCTGCTCGG 461
Db 467 GAGGCGCTATCGCCAAAGTGCCTCGCCAAAGACCGCGGACCAAGAGCGTTCTGCTCGG 526
QY 462 ATTCCGGCTCTCCGGCTCCGAGTCGTCGCGCCCGAGGGCGCGTGTGGCGGACGAGCG 521
Db 527 ATTCCGGCTCT--TGCCTCCAGTCTGTCGCGCCCGAGGGCGCGTGTGGCGGACGAGCG 584
QY 522 -CCGCGAGTCAAGGCCATCGACA-CCGCGCGGCTCCCGCGCGTGTGCTCCGG 570
Db 585 CCGCGGGGTGAAGCTATGACACCCGTCGNGNCGCGGCTGTGCTTCGG 635
```

```

RESULT 8
US-10-042-894A-18
; Sequence 18, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafaleki, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: 2002-01-09
; PRIOR FILING DATE: 2002-01-09
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(519)
; OTHER INFORMATION: n = A, T, C or G
US-10-042-894A-19

Query Match      14.3%; Score 191.6; DB 14; Length 353;
Best Local Similarity 90.2%; Pred. No. 1.6e-41;
Matches 238; Conservative 0; Mismatches 21; Indels 5; Gaps 3;

QY 667 CTCAAGCGGTGGTTCGAGGAGCAGACTCTGTTCACATCTTACTCGCGGTCGATCTCTG 726
```

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(519)
; OTHER INFORMATION: n = A, T, C, or G
US-10-042-894A-18

Query Match      22.2%; Score 298.4; DB 14; Length 519;
Best Local Similarity 89.4%; Pred. No. 3.2e-70;
Matches 328; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY 627 GTACGGAGAAAGGTGGAGTCTTGTCAAGCTGCGGAGCTCAAGGCGTGGTTCAGGA 686
Db 2 GTACGGAGAAAGGTGGAGTCTTGTCAAGCTGCGGAGCTCAANGCGTGGTTCAGGG 61
QY 687 GCAGACTGTTCCTACTCTACTCGGCGTCGATTCTTCTGGGCTATGATGCTGCTGCACT 746
Db 62 GCAGACTGTTCCTACTCTACTCGGCGTCGATTCTTCTGGGCTATGATGCTGCTGCACT 121
QY 747 CGCAGCAGCGGAGATGGGGTGGGTGACGGTGAAGCTGGTGGACTTTTGGCCATGTGCG 806
Db 122 CGCAGCAGCGGAGTGGGGTGGGTGAACAGTGAAGCTGGTGGACTTTTGGCCATGTGCG 181
QY 807 CGAGGCTGATGGGTGATTGACCAAACTTCTGGGCGGCTCTGCTCGCTGATCAAGTT 866
Db 182 CGAGGCTGATGGGTGATTGACCAAACTTCTGGGCGGCTCTGCTGATCAAGTT 241
QY 867 CTTTCTGACATGTTCCGAGACTCTCTCATAGCAGCGCTTTGGGTCTCTTAAAGAGAG 926
Db 242 TGTTCCTGACATGTTTCCAGAGACTCTCTCAGACGAGCGCTTTGGGTCTCTTAAAGAA 301
QY 927 GATCTCTGGA-TTTTCGATTGTATAACAAAGCCCTACAAGTTTGTCTGGAAGAGAGCG 985
Db 302 GATCTCTGGA-TTTTCGATTGTATAACAAAGGAGAACCTTTCAGCTGCCAAAGAGAGAG 361
QY 986 CTTCTCGA 992
Db 362 CCAAGTGA 368
```

```

RESULT 9
US-10-042-894A-19
; Sequence 19, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafaleki, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 353
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(353)
; OTHER INFORMATION: n = A, T, C, or G
US-10-042-894A-19

Query Match      14.3%; Score 191.6; DB 14; Length 353;
Best Local Similarity 90.2%; Pred. No. 1.6e-41;
Matches 238; Conservative 0; Mismatches 21; Indels 5; Gaps 3;

QY 667 CTCAAGCGGTGGTTCGAGGAGCAGACTCTGTTCACATCTTACTCGCGGTCGATCTCTG 726
```

```
Db      1  CTCAAGCATGGTGGAGGACGAGACTGTGTTCCATTCTACTCGGGCTGATCTTCTG 60
QY      727  GGCTATGATGCTGCTGAGTCGAGCGAGAGCGGGGGTGGGGTGGAGCTG 786
Db      61  GGCTATGATGCTGCTGAGTCGAGCGAGAGCGGGGGTGGGGTGGAGCTG 120
QY      787  GTGGACTTTGCCATGTGGCCGAGGGGTGATGGGT--GATTGACCACTTCTCTGGGG 844
Db      121  GTGGACTTTGCCATGTGGCCGAGGGGTGATGGGTGATTGACCACTTCTCTGGGG 180
QY      845  GGCTCTGCTGCTGATCAAGTTC--GTTTCTGACATGTTCCGGAGACTCTCATAGC-C 901
Db      181  AGCTCTGCTGATGATCAAGTTCGTTTCTTGACATGTTTCCAGANACTCTTTAGAGGC 240
QY      902  AGCCTTTGGTCTCTTCTTAAGAGA 925
Db      241  AGCCTTTGGTCTCTTCTTAANA 264

RESULT 10
US-10-042-894A-11
; Sequence 11, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1195
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (116)...(1048)
US-10-042-894A-11

Query Match
Best Local Similarity 13.7%; Score 184.6; DB 14; Length 1195;
Matches 471; Conservative 0; Mismatches 349; Indels 27; Gaps 5;

QY      56  CGCACTCCACCGCGGAGACCAAGTCGCGGCCACCGCGCTCCGCGAGCAAGCTG 115
Db      114  CCATGCTCAAGGTCGCGGATCATCAAGTCGCGGTCACCGGGGAGACGGGAAAGCTG 173
QY      116  GCCCGTCATCGAGGTCGCGCTCTTCTAAGCGGCTCCAGCGCGGGGACCGTGGG 175
Db      174  GGCACTGGTGGATGATTCGGGCGCTCTTATAAGCTCTCCAGAGCATCATCGGGAG 233
QY      176  AGCAGAGGTCGCTCTATAGGGCTTCTCCGCCACCGCGCGCTCCCGCCGATCC 235
Db      234  ACAGGAAGTGGCTTTTACGATCATTTATTTCAATACCGAGATCCAGGTCATTC 293
QY      236  GAGACACTTTCTTCCCGGTTTCCAGGACGCGACTCTCTCCCAACCGAGCGCAGCCG 295
Db      294  G---CAAAATTTCTTCTGCTTTACGGAACCTAAGACTAT---TGAGGCTCTGATGAT 347
QY      296  GGGAGCGGCACTCTCACTCTCTCGTCTCGACACTCTCTCGGGGGTTTCAGCGCCCTGG 355
Db      348  CGGGTCTCAACCTCACTCTCTCTGAGGATCTCTGTCGGGTGCGCAGAACCATCTC 407
QY      356  TCGCAGACATCAAGATCGGGCGCCATCACGTGGGCCACGAGTTCCGCGGAGCCCTACATCG 415
```

```
Db      408  TCATGACATCAAGACTCGATCCAGAACATGATATCCGAGGCGCTCTGAGGAGTACATCC 467
QY      416  CCAAGTGCCTCGCAAGGACCGCGGACCAAGAGCGTTCTGCTCGGATTCGCGGTCTCCG 475
Db      468  AANAAGTCTTAGAANAAGATCGAATAGACACAGCGTTTCATTTGGGTTTGAAGATTCTG 527
QY      476  GGCTCGAGTCTGTCGG--CCCGAGGGCGCGGTGTGGCGGACGAGCGCCCGAGGTGA 532
Db      528  GCGTAAGGGTATATCAAAATAGCGAGCTGGATTTTGGCAACCTGAGAAGAAGGTTGTT 587
QY      533  AGGCCATGACACCGCGCGCTCGCGGCTGCTCGGCGCTACGTGTCTATCC----- 585
Db      588  ATAGCTTTAATGCGGACCGGTGTCAGGTCTGAGGAAGTTGTTTCTTCCAACTTGT 647
QY      586  -----GTTCCGACAGGGGATGAGTGTGCGCTCGCGCGGCGGTGACGAGGAGAAAG 640
Db      648  CTCGGGTCCAAATGTGGATCCGATTGTTTATGATCAAAAGTTTACTGTACCGGG 707
QY      641  GTGGAGTCTTGTACAGCTGCGGAGCTCAAGCGGTGGTTCGAGGAGCAGACTCTCTTCC 700
Db      708  GTGGATTTTGGCAAAATGCTTCACTGAAGGAATGGTTTGAGGTTTCAGACGAATATC 767
QY      701  ACTTCTACTCGCGTCTGATTTCTTCTGGGTATGATGCTGTCAGTCGAGCAGCAGCGGAG 760
Db      768  ACTTCTATTCTTGTTCACCTATATCTATATGACAGGAGTCTGCTTTGGACGGC---- 823
QY      761  ATGGGGTGGGTGAGCGTGAAGCTGTGAGACTTTTCCCATGTGGCCGAGGTGATGGGG 820
Db      824  --TGTGCACACCGGAAAGTTAACTGTTGAGCTTTTGCACATGTGATGATGGCCACGG 881
QY      821  TGAATTGACACAACTTCTCGGGGCTCTGCTCGTGTATCAAGTTCGTTTCTGACATTG 880
Db      882  TGATCGATCAAACTTCTTGGTGGCTCTGTTCTGTAATCAAGTTTATACGTGACATTG 941
QY      881  TTCCGGA 887
Db      942  CTGATGA 948

RESULT 11
US-10-042-894A-9
; Sequence 9, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1105
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)...(851)
US-10-042-894A-9

Query Match
Best Local Similarity 13.3%; Score 179; DB 14; Length 1105;
Matches 460; Conservative 0; Mismatches 340; Indels 27; Gaps 5;
```

Db 15 CTCAAGATCCCGAGACACGAGTGGCGGCAAGGCCCAAGGACGAATCTGGGCCCA 74  
QY 121 CTCATCGAGCGCTCCGGCTCTTCTAAGACCGCTCCAGCGCGCCACCGTGGGAGCAC 180  
Db 75 CTCGTGACGATTTGGAAATCTACAGCCCTCCAGCCCAAGAACAGACGACACC 134  
QY 181 GAGGTGCGCTTCTATGAGCGTT-----CTCGCCCAAGCGCGCTCCCGGCCGATC 234  
Db 135 CGCGGCTCACCAACTCTCTCTTTACACCTCTCTCGCGCGCCCGCCACGACTACTCC 194  
QY 235 CGAGACACCTTCTCCCGCGTTCCAGCGACCGGACTCTCCCAACGAGCGGACCC 294  
Db 195 ATCCGCTCTCTTCTCCCGCGTTCCAGCGACCGGCTCTCTGAGCCCTCCGAGCGCTCC 254  
QY 295 GGGAGCGCGATCCTCACCTCGTCTCGAGACCTCTCGCGGGTTTCAGGGCGCCCTGC 354  
Db 255 GG---TCCCAACCTCACCTGCTCTGGAGGACTCTCTCTCGGCTACTCCAAACCTCC 311  
QY 355 GTCCGACATCAAGATCGCGGCTATCAGTGCCACCGAGTTCGCGGAGCGGCTTACATC 414  
Db 312 GTCATGACGCTAAAGATCGCTCCAGAACCTGGCACCTGGAGACTCCGAGGACTACATC 371  
QY 415 GCCAAGTGCCTCGCCAGGACCGCGGACCAAGAGGTTCTGCTCGGATTCGCGCTCTCC 474  
Db 372 TGAAGTGCCTGAAGAAGGACAGAGTCTCTAGCTTGGCTTGGTTTCAGAACTCTCG 431  
QY 475 GCGTCCGAGTCTCGCGCCCGGAGGCGCGCTGTGGCGACGAGCGCGCGGAGGTGAAG 534  
Db 432 GGAGTCAAGGACTATCTCTCTCTCTGGAACTTACCAGGAATCTCTCCAGTGTCTATCC 491  
QY 535 GCATGACACCGCGCGCTCCCGCGGTCTCGGGGT---ACGTGTCATCCGTTGCC 591  
Db 492 GCCATGCGTTCAGAGTCTGTTCTCAACAGTTCGTTCTCTTAATATCAACCATGAT 551  
QY 592 GAGCGGGATGACTGTGCTCGCGCGCGGTGTACGAGGAGAAAGTGGAGTCTTG 651  
Db 552 GATCATCATCCGATTCGCTTTCGAAACGAGGTCTAC-----GGCGCGTTTG 602  
QY 652 TCACAGTCCGCGAGCTCAAGCGTGTTCAGAGACAGACTCTGTTCACATTTCTACTCG 711  
Db 603 GAGCGTTCAGAGACTCAAGGACTGGTTCAGAGTTCAGAGCTGATCATTTCTATTCT 662  
QY 712 GCGTCAATTTCTTGGCTATGATGCTGCTGCACTGCGACGCGGAGATGGGGTGG 771  
Db 663 TGTCTCTCTTGTGGGTACGAGAGGATCTAGGAAA-----GGGAAAGCTACCAAC 716  
QY 772 GTACAGGTGAAGTGTGTGACTTTGCCATGTGGCGAGGGTGTATGGGTGATGACCAAC 831  
Db 717 CCTCTGGTCAAACTCGTTGACTTTGCACACGTGTGTGACGGAACGGTGTCTATGATC 776  
QY 832 AACTTCTGGCGGCTCTCTCGCTGATCAAGTGTGTTCTGACAT 878  
Db 777 AACTTCTGGGTGGCTTTTGTCTTCAATCAAGTTCTCAAGGATAT 823

## RESULT 12

US-10-424-599-106467  
; Sequence 106467, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 106467  
; LENGTH: 1345

; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_67156C.1  
US-10-424-599-106467

Query Match 13.3%; Score 179; DB 12; Length 1345;

Best Local Similarity 55.6%; Pred.No.5.3e-38;  
Matches 460; Conservative 0; Mismatches 340; Indels 27; Gaps 5;

QY 61 CTCACCCCGCGGAGCACCAAGTCCGCGGCCACCGCGCTCCGCCACGAGCAAGTGGGCCCG 120  
Db 220 CTCAGATCCCGGAGCACCAAGTCCGCGGCCACAGGCCAAGGACGGAATCTGGGCCCA 279  
QY 121 CTCATCGAGCGTCCGGCTCTTCTTACAAAGCGCTTCCAGGCCGGGACCGTGGGAGGAC 180  
Db 280 CTCGTGACGATTTTGGAAATTTCTACAGCCCTTCCAGACCAACAAAGACGACGACCC 339  
QY 181 GAGGTGCGCTTCTATGAGCGTT-----CTCGCCCAAGCGCGCTCCCGGCCGATC 234  
Db 340 CGCGGCTCCACGAATCTCTTTTACACTCTCTCGCGCGCGCCGCCACGACTACTCC 399  
QY 235 CGAGACACCTTCTCCCGCGTTCCACGGCACCGGACTCTCTCCCAACCGAGCGGACGCC 294  
Db 400 ATCCGCTCTCTTCCCGCGCTTTCAGGGCACCGCGCTCTCTGAGCGCTCCGAGCGCTCC 459  
QY 295 GGGAGCGCGATCTCTCACTCTCTCTCGAGACCTCTCTCGCGGGTTTCAGGGCGCTCG 354  
Db 460 GG---TCCCAACCTCACCTGCTCTGGAGGACCTCTCTCGGCTACTTCCAAACCTCC 516  
QY 355 GTCCGACATCAAGATGGCGCCATCAGTGGCCACCGAGTTCGCGGAGCCCTTACATC 414  
Db 517 GTCATGAGATGAAGATCGGCTCCAGAACCTGGACCTGGAGACTCCGAGGACTACATC 576  
QY 415 GCCAAGTGCCTCGCCAAAGACCGCGGACCAAGAGGTTCTCTCGGATTCGCGCTCTCC 474  
Db 577 TCAAGTGCCTGAAGAGGACAGAGTCTCTAGCTTGGCTTGGTTTCAGAAATCTCG 636  
QY 475 GCGTCCGAGTCTGCGGCCCGAGGCGCGTGTGGCGACGAGGCGCGCGGAGGTGAAG 534  
Db 637 GGAGTCAAGGACTCTATCTCTCTCTGGAAACCTACGAGGAAATCTCTCCAGTGTCTATCC 696  
QY 535 GCATGACACCGCGCGCTCCCGCGTCCCGGCT---ACGTGTCATCCGTTGCC 591  
Db 697 GCGCATGCTGACACTTGTCTCAACAGTTCGTTCTCTAATATATCAACATGAT 756  
QY 592 GACGAGGAGATGAGACTGTGCTCGCTCGCGCGCGTGTACGAGGAGAAAGTGGAGTCTTG 651  
Db 757 GATCATCATCCGATTTGCGCTTTCGCAACGAGGTCTAC-----GGCGCGTTTG 807  
QY 652 TCACAGTGGCGAGCTCAAGCGGTGTTCGAGGAGCAGACTCTGTTCACATTTCTACTCG 711  
Db 808 GAGCGCTTCGAGAGCTCAAGGACTGTTTCGAGGTTTCAGCGGTGTATCATTTCTATTCT 867  
QY 712 GCGTCAATTTCTGGGCTATGATGCTGCTGAGTTCGAGCTCGACGAGCGGAGATGGGGTGG 771  
Db 868 TGTCTGTTCTTGTGGTGTACGAGAGGATCTAGGAAA-----GGGAAAGCTACCAAC 921  
QY 772 GTGACGCTGAAGCTGGTGGACTTTGCCATGTGGCGCGGAGGTGTATGGGTGATGACCAAC 831  
Db 922 CCTCTGGTCAAACTCGTTGACTTTGCACACGTGTGTGACGGAACCGGTGTCTATTGATC 981  
QY 832 AACTTCTGGCGGCTCTGCTCGCTGATCAAGTGTGTTCTGACAT 878  
Db 982 AACTTCTGGGTGGCTTTGTTCTCTTCAATCAAGTTCTCAAGGATAT 1028

## RESULT 13

US-10-042-894A-13  
; Sequence 13, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui

```
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011alel Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042, 894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Parthenium argentatum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21)...(908)
; US-10-042-894A-13

Query Match 11.1%; Score 149.8; DB 14; Length 1020;
Best Local Similarity 53.2%; Pred. No. 3.5e-30;
Matches 456; Conservative 0; Mismatches 362; Indels 39; Gaps 5;

QY 61 CTCACCCCGCGGAGCACCAAGTCGCGGCCACCGCGCCTCCGCCAGCAGCTGGGCGCG 120
DB 24 CTCGAGGCCCGAGATCATCAGGTTGCTGGACATGAAGCTGGGCTCGGAGCTTGGCCCA 83
QY 121 CTATCGACGGCTCGCGCTCTTACAGCGCTCCAGCGCGCGCGCGCGCGCGGAGCAC 180
DB 84 CTATTGATGATTGAGCGCGGTTTACAAACCATGTCAGGGTGATACCCGTTGGGTGAGAA 143
QY 181 GAGGTGCGCTTCTATGAGCGGTTCTCCGCCAGCGCGCGCTCCGCCCGCCATCCGAGAC 240
DB 144 GAAGTAGCCTTTTATGATCATTTCTTCTTAAATAATATTCAGAACACATACG--C 200
QY 241 ACCTTCTTCCCGGTTCCAGGCGCGAGCTCTCTCCCAACCGAGGCGAGCCCGGGAG 300
DB 201 AAATTTCTTCTATATATTATGACCAACAAAT-----CATGAAGGCATCCACTGGCTCT 254
QY 301 CCGATCTCTCACTGCTCTCGAGACCTCTCGGGGTTTCAGGGCGCTCGGTCCGA 360
DB 255 GACATCTCTCACATGGTGTTCAGAGATCTTACATGAGCTCATGTCAACCCATCTGTAATG 314
QY 361 GACATCAAGATCGCGGCATCAGTGGCCACCGAGTTTCGCGGAGCCCTACATCGCCAAAG 420
DB 315 GACATCAAAATCGGGTCCAGAACATGGCGCCAGAGCTTCGAGGGGTACATTCGAAA 374
QY 421 TGCTCGCCAAAGACCGCGGAGCACGAGGTTCTGTCTCGGATTCGCGGTCCTCGGGGTC 480
DB 375 TGCTTAAAAAAGGATAGGGAAGCACAAAGTATTCATTGGGATTCAGGATCTCGGGGCTG 434
QY 481 CGAGTCTGCGCCCGGAGGCGCGTGTGGCGGACGG--AGCGCCCGAGGTTGAAGGCC 537
DB 435 CAAGTCTATATGATGATGGTTCAGGGTTTTTAAGCCTCATAGAAATTCATGCGTAA 494
QY 538 ATGGACACCGCGCGCTCCGCGGTCTCGCGGCTACGTTGTCATCCGTTGCGAGCAG 597
DB 495 ACCGCGCCAGCTGATGTTAGACTACTTCTTAGGAAATTTGTTCTTCTAACCCTCTGCA 554
QY 598 GGGAT-----GGACTGTGCGCTTCGCGCGCGCGCGGCGGTGACGGA 633
DB 555 GAGATGGAATGCGCACAGGCCTAGGCGCGGATTTCTTTAGCATCTTTTGTATGGT 614
QY 634 GGAAGGTTGAGTCTTGTACAGCTCGGAGCTCAAGGGTGTTCGAGGAGCAGACT 693
DB 615 GGGCCTAATGGGATATTAGCTCAACTGATGGAATGGAAGACATGGTTTGAAGATCAACA 674
QY 694 CTGTTCACATCTACTCGGCGTCAATTTCTTGGGCTATGATGTGTCGACGTGCGAGCA 753
DB 675 ATTTACCACATCTATGCTGTTCTTTTGTTCATCTTTGAAAGAGGTTGGTGTAA 734
```

```
QY 754 GCGGAGATGGGGTGGGTGACGGTGAAGCTGGTGGACTTTTCCCATGTGGCGGAGCT 813
DB 735 GGT---GCTCGGTCAAAACGACAGAGTCAAACTTAATGATTTTGTCTCATGTACAGATGGT 791
QY 814 GATGGGTCATTGACCAACAACTTCCCTGGCGGGCTCTGCTCGCTGATCAAGTTCGTTTCT 873
DB 792 AATGGTGTATTGATCACAATTTCTTGGGTGGGCTCTGTTCTTTGATAAAGTTCATTTCT 851
QY 874 GACATTGTTCCGGAGAC 890
DB 852 GACATACTTTCGGAGAC 868
```

## RESULT 14

```
US-09-770-444-340/c
; Sequence 340, Application US/09770444
; Patent No. US20020023280A1
```

## GENERAL INFORMATION:

```
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2027 (PARA-016PRV)
; CURRENT APPLICATION NUMBER: US/09/770,444
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,502
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 340
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-770-444-340
```

```
Query Match 7.8%; Score 104.6; DB 9; Length 464;
Best Local Similarity 64.2%; Pred. No. 4.2e-18;
Matches 174; Conservative 0; Mismatches 94; Indels 3; Gaps 1;
```

```
QY 618 CGCGCGGTGTACGGAGGAAAAGTGGAGTCTTGTGCACAGCTCGCGAGTCAAGGGTG 677
DB 464 CTCGAGTGTTTAGCGGGTTCACCGGGATCTTAACGCGAGTTGCTGGAACCTCAAGACCTG 405
QY 678 GTTCGAGGACGACACTCTGTCCACTTCTACTCGGGCTGATTTCTTGGGCTATGATGC 737
DB 404 GTTCGAGAACCAACACCTCTACCATTTCAACTCTTGTTCGATTTTAAATGTCATAGAA 345
QY 738 TGCTGCAGTCGCA---GCAGGCGGAGATGGGGTGGGGTGACGGTGAAGCTGGTGGACTT 794
DB 344 TGAATCACTTAAAGGAAATGATGATGCTAGACCACAAAGTCAAGTCTGATGGATTT 285
QY 795 TGCCCATGTGGCGAGGGTGAATGGGTGATGACCAACTTCTCTGGCGGGGCTCTGCTC 854
DB 284 TGCTCATGTTCTTGTGATGGTAAATGGTGTCAATGACCACTACTTCTTGGGTGGTCTTCTGCTC 225
```

QY 855 GCTGATCAAGTTCGTTTCTGACATTTGTTCCG 885  
Db 224 TTTCATAAACTTCATTGTCGATTTCTTCAG 194

RESULT 15  
US-10-314-657-1/c  
; Sequence 1, Application US/10314657  
; Publication No. US20030175888A1  
; GENERAL INFORMATION:  
; APPLICANT: SHEN, Ben  
; APPLICANT: CHENG, Yi-Qiang  
; APPLICANT: TANG, Gong-Li  
; TITLE OF INVENTION: Discrete Acyltransferases Associated with Type I Polyketide  
; TITLE OF INVENTION: Synthases and Methods of Use  
; FILE REFERENCE: 054030-0021  
; CURRENT APPLICATION NUMBER: US/10/314,657  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: PCT/US02/08937  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: US 60/278,935  
; PRIOR FILING DATE: 2001-03-26  
; NUMBER OF SEQ ID NOS: 214  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 135638  
; TYPE: DNA  
; ORGANISM: Streptomyces atroolivaceus  
US-10-314-657-1

Query Match 5.3%; Score 71.4; DB 14; Length 135638;  
Best Local Similarity 43.9%; Pred. No. 1.4e-08;  
Matches 306; Conservative 0; Mismatches 391; Indels 0; Gaps 0;

QY 59 ACCTCCACCCGCGGAGACCAAGTCGCGGCCACCGCGCTCCGCCAGCAAGCTGGGCC 118  
Db 40684 AGCGCGCGTGGCAGAGCTCAGCCCCGAGCGCGCCCTGTCTGCGCCGCCACCTGGCCCC 40625  
QY 119 CGCTCATCGACGGCTCCGGCCCTCTTCTAAGCCCGCTCCAGCCGCGGACCGTGGGGAGC 178  
Db 40624 GCGGAGGGCGGCTCGGCATCTCGGTACAGGCTCTGCCGCGCACCGGGGAGGCACAGC 40565  
QY 179 ACGAGGTGCGCTTCTATAGGCGTTCTCCGCCACGCGCGTCCCGCGCCGCGCATCCGAG 238  
Db 40564 GCTACGTGGCATCCCGGCGCCAGCAGAGCTCTACTACAGTACAGCTCGACCCGCTCG 40505  
QY 239 ACACCTTTCTCCCGGGTTCCAGGCGAGGACTCTCCCGACCGAGGCGGAGCCCGGGG 298  
Db 40504 CGGTGAACCTACCTGTGCCGATCGCGGTGCGGCTGCGGCGGACCTGGACACCCCAAGGC 40445  
QY 299 AGCGCATCTCTACCTCGTCTCGACACCTCTCTCGGGGGTTTCAGGCGCCCTGCGTCG 358  
Db 40444 TGGAGCGGGCCCTGCGCGGCTGCTGCCCGCACAGGCGCTGGGCACCCGCTTCGGA 40385  
QY 359 CAGACATCAAGATCGGCGCCATACAGTGGCCACCGAGTTCCCGGAGCCCTACATCGCCA 418  
Db 40384 CCGACCCCGAACACGGCGTGTGCGAGTCTGCGCACCCCGACAGGTGCGCTCGCAC 40325  
QY 419 AGTGCTTCGCCAAGACCGGGACCGAGCGTTCTGCTCGGATTCGCGCTCTCCGCG 478  
Db 40324 TCGGACGCGCGCGGCGCGCGCGCGGAGGTGCGCGCACTCTCTGACCGGATCGCCCGGC 40265  
QY 479 TCCGAGTCTCGGCCCGAGGCGCGCTGTGTGGCGAGCGAGCGCCCGAGGTGAAGCCA 538  
Db 40264 GCCCTTCGACCTGTGCGCGCGCGCTGTGGCGGCAAGCTGTGGCAGCTCACCGACG 40205  
QY 539 TGGACACCGCGCGCTCGCGCGCTGTCCGCGCTACGTGTATCCGTTGCCACGAGG 598  
Db 40204 ACCCGAGGCGCGCGCGCGCTCTCTGCGTGTGCTCACCTGTGCTCCACCATCTGTCG 40145  
QY 599 GGATCGACTGTGCGCTCGCGCGGTGTACGGAGGAAAGGTGGAGTCTTGTACAGC 658  
Db 40144 TGGACGCGTGTGCTCGCGGTGCTCTCGAGAACTCACGAGGGGTACGCGGCCGAGC 40085

QY 659 TCGCGGAGCTCAAGCGTGGTTTCGAGGAGCAGACTCTGTTCACACTTCTACTCGGCGTCGA 718  
Db 40084 TCACCGGGGCGCGCGGAACCTGCCCCGAACCTGCGCGTGCAGTACCGGACTTCGCGCGCG 40025  
QY 719 TTCTTCTGGGCTATGATGCTGTGCTGCAAGTGCAGAGG 755  
Db 40024 CCCATCACGATGGCTGGCGGTTGACGAGGCGGCACG 39988

Search completed: March 27, 2004, 06:11:15  
Job time : 543 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 20:55:34 ; Search time 20 Seconds  
(without alignments)  
1389.968 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521  
Sequence: 1 MSDLHPHQVAGHRASAK.....IKFVSDIVPETHQPLGPS 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	673.5	44.3	286	2	T49876
2	207.5	13.6	323	2	T21569
3	183	12.0	351	2	T17246
4	167.5	11.0	268	2	T50224
5	146.5	9.6	242	2	T27999
6	145	9.5	461	2	JN0129
7	131	8.6	459	2	S13064
8	123	8.1	355	1	RGBYR3
9	118	7.8	1050	2	S54640
10	116.5	7.7	967	2	T41672
11	114.5	7.5	673	2	S41053
12	110	7.2	394	2	T42512
13	110	7.2	461	2	T42513
14	110	7.2	494	2	T42444
15	109.5	7.2	472	2	S17682
16	109.5	7.2	946	2	JC7810
17	103.5	6.8	280	2	T25898
18	98	6.4	1546	2	F75461
19	97.5	6.4	332	2	S73031
20	97.5	6.4	362	2	E86996
21	93.5	6.1	756	2	F70815
22	90.5	6.0	486	2	T25639
23	90.5	6.0	1678	2	T35547
24	90	5.9	1155	2	G87477
25	89	5.9	380	2	C83719
26	89	5.9	4613	2	T17409
27	88.5	5.8	370	1	B43306
28	88.5	5.8	377	1	BOAG58
29	88.5	5.8	377	2	AF3249

30	88.5	5.8	867	2	T14777
31	88	5.8	327	2	D83320
32	87.5	5.8	407	2	T30585
33	87.5	5.8	409	2	F75261
34	87.5	5.8	550	2	T10393
35	87	5.7	363	2	H87464
36	87	5.7	391	2	T08338
37	87	5.7	753	1	D72660
38	86.5	5.7	313	2	D87336
39	86.5	5.7	373	2	F70781
40	86.5	5.7	409	1	BOAG55
41	86.5	5.7	454	2	T47253
42	86.5	5.7	1032	2	P65071
43	86	5.7	605	2	T35277
44	85.5	5.6	319	2	D75349
45	85.5	5.6	337	2	B84335

#### ALIGNMENTS

##### RESULT 1

T49876

hypothetical protein T211.80 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000

C/Accession: T49876

R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; L

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24493

A:Accession: T49876

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-286 <BEV>

A:Cross-references: EMBL:AL163912; GSPDB:GN00063; ATSP:T211.80

A:Experimental source: cultivar Columbia; BAC clone T211

C:Genetics:

A:Gene: ATSP:T211.80

A:Map position: 5

Query Match 44.3%; Score 673.5; DB 2; Length 286;

Best Local Similarity 50.4%; Pred. No. 4.1e-48;

Matches 141; Conservative 42; Mismatches 86; Indels 11; Gaps 7;

Qy	4	LHPPEHQVAGHRASAKLGLIDSGLFYKPLQAGDGEHEVAYEAFSAHAAYPARIRD	63
Db	3	LKVPHEQVAGHIAKDGPGLVDKGRFFKPLQGDGRGEIEVKFYESFSSTVEPEHLH-	61
Qy	64	TFPFRFHGTRLLPTEAQPGEHPHLVLDLILAGFQAPCVADIKIGAITWPPSPPEPYIAK	123
Db	62	RYFPVYHGTQAV--EGSDGA--AMVLENLLAEYTKSVMDVKGSRTWYDPDASEEYIQK	117
Qy	124	CLAKDRGTTSVLLGFRVSGVRVVG--PGAVVTRTERPEVKAMDTAGVRRLRYVYS--SVA	180
Db	118	CLKXDTGTVTSSGFRISGFEVYDHKESSEFWKPKRLIRGLDVGDCARLTLRKFVSSNSLS	177
Qy	181	DEGM--DCALAAAYGKGKGVLSQLRELKAWFEOTLFHFYSASILLGYDAAVAAGDGG	238
Db	178	DTGSKPDASAFSSVGGSHGILLTOLLELKTWFENQTLYHFNSCSILMYENESILKGNDD	237
Qy	239	GG-VTVKLVDFAHVAEGDGVGDHNFGLGCLSLIKFVSDIV	277
Db	238	DARQVKLVDFAHVLDGNGVIDHNHFLGLGCLSFNFIREIL	277

##### RESULT 2

T21569

hypothetical protein F30A10.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Sep-2000

C/Accession: T21569

R:Barlow, K.

submitted to the EMBL Data Library, October 1996

hypothetical prote	
hypothetical prote	
transferase homolo	
cytochrome P450 -	
chitinase (EC 3.2.	
nitrogen regulatio	
conserved hypothet	
probable aldehyde	
riboflavin biosynt	
probable citA prot	
virB10 protein - A	
core protein II (i	
hypothetical prote	
probable transfera	
conserved hypothet	
hypothetical prote	

QY	279	ET 280
Db	348	ES 349
<p>RESULT 4</p> <p>T50224</p> <p>probable arginine metabolism transcription control protein [imported] - fission yeast</p> <p>C;Species: Schizosaccharomyces pombe</p> <p>C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000</p> <p>C;Accession: T50224</p> <p>R;Zimmermann, W.; Wambutt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G. submitted to the EMBL Data Library, December 1999</p> <p>A;Reference number: Z25047</p> <p>A;Accession: T50224</p> <p>A;Status: preliminary; translated from GB/EMBL/DBJ</p> <p>A;Molecule type: DNA</p> <p>A;Residues: 1-268 &lt;ZIM&gt;</p> <p>A;Cross-references: EMBL:AL1135751; PIDN:CAB63791.1; GSPDB:GN00066; SPDB:SPAC607.04</p> <p>A;Experimental source: strain 972h(-); cosmid c607</p> <p>C;Genetics:</p> <p>A;Gene: SPDB:SPAC607.04</p> <p>A;Map position: 1</p>		
QY	88	LVLLDLAGFOAPCVADIKIGAITWPPSPPEPIAKCLAKDRGTTSVLLGFRVSGVRVVG 147
Db	83	IILNLYIQMETPCVMDIKLGKQLWADDAPLEKRRKLDVSRSTTSGSLGFRITGI---- 138
QY	148	PEGAVW-RTPEVKAMDTAGVRRVLRVYSSVADEGMDCALAAAYGKGVLSQLREL 206
Db	139	---LSWDRNTNTVIKSTAWGKTLTDSVVVEGLNDFVSCSLSQ-----KARLVESFLNL 190
QY	207	KAWFE---RQTLPHFYSAILLGYDAAV-AAGDGGGVTVKLVDFAHVAEGDGVIDNF 262
Db	191	LKLFVDLSSESYIELKSSILFVYDYSNLNPTVHCSNVVLKLIIDIAHSRWTKNTDHTNT 250
QY	263	LGGLCSLI 270
Db	251	LIGVKMLI 258
<p>RESULT 5</p> <p>T27999</p> <p>hypothetical protein ZK795.1 - Caenorhabditis elegans</p> <p>C;Species: Caenorhabditis elegans</p> <p>C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000</p> <p>C;Accession: T27999</p> <p>R;Percy, C.</p> <p>submitted to the EMBL Data Library, December 1996</p> <p>A;Reference number: Z20453</p> <p>A;Accession: T27999</p> <p>A;Status: preliminary; translated from GB/EMBL/DBJ</p> <p>A;Molecule type: DNA</p> <p>A;Residues: 1-242 &lt;WIL&gt;</p> <p>A;Cross-references: EMBL:Z83246; PIDN:CAB05842.1; GSPDB:GN00022; CESP:ZK795.1</p> <p>A;Experimental source: clone ZK795</p> <p>C;Genetics:</p> <p>A;Gene: CESP:ZK795.1</p> <p>A;Map position: 4</p> <p>A;Introns: 36/1; 59/3; 88/2; 116/1</p> <p>C;Superfamily: Caenorhabditis elegans hypothetical protein ZK795.1</p>		
QY	10	QVAGHRASAKLGP-----LIDGSGLFYKPLQAGRGHEVAFYAFSAHAAPFA----- 59
Db	19	OTAGHPSPVINGKREIGLTKIPGSRILKPKDASRGKEVALYELLSCSTTSPSTPPE 78



## RESULT 8

RGYR3  
regulatory protein ARG82 - yeast (Saccharomyces cerevisiae)  
N/Alternate names: protein YD9335.06c; protein YDR173c; regulatory protein ARGRIII  
C/Species: Saccharomyces cerevisiae  
C/Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jun-2000  
C/Accession: S05823; S49769  
R/Dubois, E.; Bercy, J.; Messenguy, F.  
Mol. Gen. Genet. 207, 142-148, 1987  
A/Title: Characterization of two genes, ARGRI and ARGRIII required for specific regulation of arginine biosynthesis in *Saccharomyces cerevisiae*  
A/Reference number: S05822; MUID:87257286; PMID:3298999  
A/Accession: S05823  
A/Molecule type: DNA  
A/Residues: 1-355 <DUB>  
A/Cross-references: EMBL:X05328; NID:g3375; PIDN:CAA28945.1; PID:g3376  
A/Note: the authors translated the codon AAG for residue 54 as Arg and GAG for residue 1  
R/Murphy, L.; Harris, D.E.  
submitted to the EMBL Data Library, November 1994  
A/Reference number: S49764  
A/Accession: S49769  
A/Molecule type: DNA  
A/Residues: 1-355 <NUR>  
A/Cross-references: EMBL:Z46727; NID:g1289283; PIDN:CAA86678.1; PID:g1289288; GSPDB:GN00  
C/Genetics:  
A/Genes: SGD:ARG82; ARG83; MIPS:YDR173C  
A/Cross-references: SGD:S0002580; MIPS:YDR173C  
A/Map position: 4R

C/Function:  
A/Description: responsible for repression of 6 genes and induction of 2 genes in response  
C/Superfamily: regulatory protein ARGRIII  
C/Keywords: nucleus; transcription regulation  
F:282-301/Region: aspartic acid-rich

Query Match 8.1%; Score 123; DB 1; Length 355;

Best Local Similarity 20.3%; Pred. No. 0.013;

Matches 77; Conservative 46; Mismatches 100; Indels 156; Gaps 18;

QY 8 EHOVAGHRASAKGLPLDGSQ-LFYKPLQAGDRCGHEVAFYEAFSAHAAVPAIRD--- 63  
DB 11 EHKAAGHD-----GULTDGDGLLIFKP-----APQLEFYKAI-----QVRDYSR 51  
QY 64 -----TFFRFHGT-----RL-----LPTAQPG 82  
DB 52 RKSSADGDAPLCSWMFTYGLVNEGAKIBQSGDAALLKIDRLSDSTNDLSDIPVKSEK 111  
QY 83 EHPHVLVDLLAGFOAPCVADIKGAIATWPPSPFPYIAKCLAKDRGTTSVLLGFRVSG 142  
DB 112 K-QYLVLENLLYGFSGKPNLDIKLTKLYDSKASLEKRMKRVSETTSGSLFRICG 169  
QY 143 VRVVGEGAVWRTERPEVKAMDTAGVRRVLRVSVSSVADEGMDCAAAYVGG----- 195  
DB 170 MKT-----QKNFSV-----LNQLSLEYEEADS--DYIFINKLYGRSRTDQNV 211  
QY 196 -----KGVLSQIRE---LKAMFEEQTLFH-----FYSASILLY----- 227  
DB 212 SDATIELYFNNPHLSDAKQHQLKTKLQLFYNTMLEBEVRMISSSLFIYEGDPERWE 271  
QY 228 -----DAAVAAG-----GDGGGVTKLVDFAH--V 251  
DB 272 LLNDVDKLMRDFFIDDDDDDDDDDDDAEGSGEPKDKTKTSGLSMSLIDFAHSEI 331  
QY 252 AEGDGVIDHNFGLGCLSLI 270  
DB 332 TPCKG-YDENVIEGVETLL 349

## RESULT 9

S54640  
KCS1 protein - yeast (Saccharomyces cerevisiae)  
N/Alternate names: protein D3238; protein PF1050; protein YD9335.03c; protein YDR017c  
C/Species: Saccharomyces cerevisiae  
C/Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 21-Jul-2000  
C/Accession: S54640; S63423; S67830; S60373; S72114

R/Dedman, K.; Brown, D.; Hamlyn, N.; Bowman, S.  
submitted to the EMBL Data Library, May 1995  
A/Reference number: S54638  
A/Accession: S54640  
A/Molecule type: DNA  
A/Residues: 1-1050 <DED>  
A/Cross-references: EMBL:Z49770; NID:g840867; PID:g840870  
A/Experimental source: strain AB972  
R/Eide, L.G.; Sander, C.; Prydz, H.  
submitted to the EMBL Data Library, February 1996  
A/Description: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome IV  
A/Reference number: S63416  
A/Accession: S63423  
A/Molecule type: DNA  
A/Residues: 1-1050 <EID>  
A/Cross-references: EMBL:X95966; NID:g1216215; PID:e253292; PID:g1431442; MIPS:YDR017C  
R/Prydz, H.; Eide, L.G.  
submitted to the Protein Sequence Database, July 1996  
A/Reference number: S67822  
A/Accession: S67830  
A/Molecule type: DNA  
A/Residues: 1-1050 <PRY>  
A/Cross-references: EMBL:Z74313; NID:g1431441; PID:e253292; PID:g1431442; MIPS:YDR017C  
A/Experimental source: strain S288C  
R/Huang, K.N.; Symington, L.S.  
Genetics 141, 1275-1285, 1995  
A/Title: Suppressors of a Saccharomyces cerevisiae pck1 mutation identify alleles of the  
A/Reference number: S60373; MUID:96170769; PMID:8601473  
A/Accession: S60373  
A/Status: not compared with conceptual translation  
A/Molecule type: DNA  
A/Residues: 1-688, 'M', 690-1050 <HUA>  
R/Eide, L.G.; Sander, C.; Prydz, H.  
Yeast 12, 1085-1090, 1996  
A/Title: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome IV  
A/Reference number: S72107; MUID:97051598; PMID:8896275  
A/Accession: S72114  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-1050 <EW>  
A/Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65208.1; PID:g1216223  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996  
C/Genetics:  
A/Genes: SGD:KCS1  
A/Cross-references: MIPS:YDR017C; SGD:S0002424  
A/Map position: 4R

Query Match 7.8%; Score 118; DB 2; Length 1050;

Best Local Similarity 24.4%; Pred. No. 0.13; 79; Indels 58; Gaps 9;

Matches 53; Conservative 27; Mismatches 27;

QY 88 LVLDLLAGFOAPCVADIKGAIATWPPSPPE-----PYIAKCLAKDRGTTSVLLGFRVSGV 143  
DB 760 ILLEDLTRMNKPCALDLKMGTRQYGVDAKRAKQLSQRAKCLK---TTSRLGVRICGL 815  
QY 144 RVVGEGAVWRTERPEVKAMDTAGVRRVLRVSVSSVADEGMDCA--LAAAVYGGKGVLS 201  
DB 816 KVMNKDYITRD-----KYFGRVRVKGWQFARVLARFLYDGR-TIES 856  
QY 202 QRELKAMFEE-QTL-----PHEYSASILLGYDAAVAAG--DGGGVTVKLVDF 248  
DB 857 LIHQIPRLIKQDITLYSEIFNLKGYRLYGASLLMYDGDANKSNKRKKAANKVNLIDP 916  
QY 249 AHVA-----EGDGVIDHNFGLGCLSL 269  
DB 917 ARCVTIKEDAMECMKFERIPPKSPFNIEDKGLFAGVKSLL 953

## RESULT 10

T41672  
probable b-zip transcription factor - fission yeast (Schizosaccharomyces pombe)  
C/Species: Schizosaccharomyces pombe  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

DB 346 IFGNEDGYLGIDNVKLEEL-PE 370

RESULT 13

T42513

1P-myo-inositol-trisphosphate 3-kinase (EC 2.7.1.127) isoform 3 - Caenorhabditis elegans

C-Species: caenorhabditis elegans

C-Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 21-Jul-2000

C-Accession: T42513

R-Clandinin, T.R.; DeModena, J.A.; Sternberg, P.W.

Cell 92, 523-533, 1998

A-Title: Inositol triphosphate mediates a RAS-independent response to LET-23 receptor

A:Reference number: Z22166; MUID:98150857; PMID:9491893  
A:Accession: T42513  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-461 <CLA>  
A:Cross-references: EMBL:AF045613; NID:g2898161; PIDN:AAC38962.1; PID:g2898162  
C:Genetics:  
A:Gene: LFE-2  
C:Keywords: phosphoprotein; phosphotransferase

Query Match 7.2%; Score 110; DB 2; Length 494;  
Best Local Similarity 21.8%; Pred. No. 0.23;  
Matches 71; Conservative 46; Mismatches 99; Indels 110; Gaps 18;  
  
QY 10 QVAGHRASAKGLPLIDGSLFYKPLQAGDRGEHEVAFAEYSAHAAPARIRDTFFPRF 69  
DB 166 QUSGHEGIVPATP-----HTLYKK-QCANGCEGRA--YKNISKDPAL-----DGTFKY 212  
  
QY 70 HGTRLLPTEAQGPGEHPHLVLDLLAGFQAP---CVADIKIGAITWPPS-----115  
DB 213 Y-----DELEKNEER-FIEIEDLLQGFHDPTKTAIMDKIGTRTFLESEVSNTKKRADL 265  
  
QY 116 -----SPEPIAKCLA-----KDRGTTSVLLGFRVSGVRVVGPEGA-----151  
DB 266 YERKWAINDPEETEERKCGAITKLRVMOFRRESSTQAQGFRIEAAKRL--EGALEKNF 323  
  
QY 152 --VWRTPEVKAMDTAGVRRVRLRYVSSVADGMDCALAAAVGKGVLSQLRELKAW 209  
DB 324 KKVRTVEDVTTFMDFFGTOR-----SRVROQ-----LIERLKSWMKA 361  
  
QY 210 FEEQTLPHFY---SASILLGYDAANAAGGGGGVTVKLVDFAH---VAEG-----254  
DB 362 IEHSSFFNSHEVVGSSILIVDFTEKVGCM-----MIDFAKSSPVNGRTLNRHTTW 412  
  
QY 255 -DGVIDHNFGLGCLSLIKFVSDIVPE 279  
DB 413 IPGNNEEDGYLIGIDNLVKILEEL-PE 437

RESULT 14  
T42444  
1D-myo-inositol-trisphosphate 3-kinase (EC 2.7.1.127) 1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jul-2000  
C:Accession: T42444  
Cell 92, 523-533, 1998  
A:Title: Inositol trisphosphate mediates a RAS-independent response to LET-23 receptor  
A:Reference number: Z22166; MUID:98150857; PMID:9491893  
A:Accession: T42444  
A:Molecule type: mRNA  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Residues: 1-494 <CLA>  
A:Cross-references: EMBL:AF045611; NID:g2898157; PIDN:AAC38960.1; PID:g2898158  
C:Genetics:  
A:Note: LFE-2  
C:Keywords: phosphotransferase

Query Match 7.2%; Score 110; DB 2; Length 494;  
Best Local Similarity 21.8%; Pred. No. 0.23;  
Matches 71; Conservative 46; Mismatches 99; Indels 110; Gaps 18;  
  
QY 10 QVAGHRASAKGLPLIDGSLFYKPLQAGDRGEHEVAFAEYSAHAAPARIRDTFFPRF 69  
DB 199 QUSGHEGIVPATP-----HTLYKK-QCANGCEGRA--YKNISKDPAL-----DGTFKY 245  
  
QY 70 HGTRLLPTEAQGPGEHPHLVLDLLAGFQAP---CVADIKIGAITWPPS-----115  
DB 246 Y-----DELEKNEER-FIEIEDLLQGFHDPTKTAIMDKIGTRTFLESEVSNTKKRADL 298  
  
QY 116 -----SPEPIAKCLA-----KDRGTTSVLLGFRVSGVRVVGPEGA-----151  
DB 299 YERKWAINDPEETEERKCGAITKLRVMOFRRESSTQAQGFRIEAAKRL--EGALEKNF 356

QY 152 --VWRTPEVKAMDTAGVRRVRLRYVSSVADGMDCALAAAVGKGVLSQLRELKAW 209  
DB 357 KKVRTVEDVTTFMDFFGTOR-----SRVROQ-----LIERLKSWMKA 394  
  
QY 210 FEEQTLPHFY---SASILLGYDAANAAGGGGGVTVKLVDFAH---VAEG-----254  
DB 395 IEHSSFFNSHEVVGSSILIVDFTEKVGCM-----MIDFAKSSPVNGRTLNRHTTW 445  
  
QY 255 -DGVIDHNFGLGCLSLIKFVSDIVPE 279  
DB 446 IPGNNEEDGYLIGIDNLVKILEEL-PE 470

## RESULT 15

S17682  
1D-myo-inositol-trisphosphate 3-kinase (EC 2.7.1.127) B - human  
N:Alternate names: inositol-1,4,5-trisphosphate 3-kinase B  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 21-Jul-2000  
C:Accession: S17682  
R:Takazawa, K.; Perret, J.; Dumont, J.E.; Erneux, C.  
Biochem. J. 278, 883-886, 1991  
A:Title: Molecular cloning and expression of a new putative inositol 1,4,5-trisphosphate  
A:Reference number: S17545; MUID:91378954; PMID:1654894  
A:Accession: S17682  
A:Molecule type: mRNA  
A:Residues: 1-472 <PAK>  
A:Cross-references: EMBL:X57206; NID:g333990; PIDN:CAA40491.1; PID:g333991  
C:Genetics:  
A:Gene: GDB:ITPKB  
A:Cross-references: GDB:128973; OMIM:147522  
A:Map position: lq41-lq43  
C:Keywords: phosphotransferase

Query Match 7.2%; Score 109.5; DB 2; Length 472;  
Best Local Similarity 20.3%; Pred. No. 0.24;  
Matches 55; Conservative 40; Mismatches 85; Indels 91; Gaps 13;  
  
QY 65 FFRFRHTRLLPTEAQGPGEHPHLVLDLLAGFQAPCVADIKIGAITW-----112  
DB 237 FVPAYHG-----DVVKDGERYNQ--MDDLIAFDSPCWMDCMGIITYLEELTKARKKP 289  
  
QY 113 -----PPSP--EPYIAKCLAKDQG-----TTSVLLGFRVSGVRVVGPEG 150  
DB 290 SLRKMVYQRMIEVDPEAPTEEERAKRAVTKPRYMQWRETISTATLGFRIEGIK--KEDG 347  
  
QY 151 AVWRTPEVKAMDTAGVRRVRLRYVSSVADGMDCALAAAVGKGVLSQLREL 206  
DB 348 TV---NRDFKTKTREQVTEAPREFTK-----GNHNILIAVRDLKAI 387  
  
QY 207 KAWFEEQTLPHFY---SASILLGYDAANAAGGGGGVTVKLVDFAH---VAEG-----254  
DB 388 RTTLEVSPPFKCHEVIGSSLLFIHDKKEQA-----KVMYIDFGTKTLPPEGQTLOHD 439  
  
QY 255 ----DGVIDHNFGLGCLSLIKFVSDIVPE 281  
DB 440 VPMQEGNRDGYLSGLNVLVDILTEMSQDAP 470

Search completed: March 24, 2004, 20:59:40  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 27, 2004, 06:02:11 ; Search time 96 Seconds  
(without alignments)  
1670.633 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPHQBHVAGHRASAK.....IKFVSIVPHTPTQPLGPS 289

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framer\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO spoil/US10042894/runat\_24032004\_145049\_16982/app\_query.fasta.1.455  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10042894 -CGN\_1\_69 @runat\_24032004\_145049\_16982 -NCPUS=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	193	12.7	1341	4	US-09-023-655-809
2	189	12.4	1748	4	US-09-3120-712
3	113.5	7.5	1899	4	US-09-252-991A-6525
4	107.5	7.1	1984	4	US-09-252-991A-12872
5	107.5	7.1	4078	4	US-09-016-434-1132
6	106.5	7.0	6000	1	US-08-348-006B-6
7	106.5	7.0	6000	2	US-08-800-825A-6
8	106.5	7.0	6000	3	US-09-158-657-6
9	106.5	7.0	6000	5	FCI-US94-10166-6
10	106	7.0	4411529	3	US-09-103-840A-1
11	105.5	6.9	36412	4	US-08-311-731A-132
12	104	6.8	1917	2	US-08-637-899-2

13	103.5	6.8	1896	4	US-09-252-991A-11331
14	100.5	6.6	960	4	US-09-252-991A-12565
15	100.5	6.6	2514	4	US-09-144-914-3
16	100.5	6.6	3446	4	US-09-620-312D-653
17	100	6.6	1425	4	US-09-489-039A-1509
18	100	6.6	1491	4	US-09-252-991A-6021
19	100	6.6	2016	4	US-09-489-039A-1492
20	99	6.5	1470	4	US-09-489-039A-6625
21	98	6.4	4403765	3	US-09-103-840A-2
22	98	6.4	4411529	3	US-09-103-840A-1
23	96.5	6.3	1509	4	US-09-489-039A-6892
24	96.5	6.3	4403765	3	US-09-103-840A-2
25	96	6.3	1230	4	US-09-252-991A-14780
26	96	6.3	1560	2	US-08-709-979A-2
27	96	6.3	23673	4	US-09-773-816-1
28	95	6.2	813	4	US-09-489-039A-5090
29	95	6.2	1952	4	US-09-566-921-24
30	95	6.2	2154	4	US-09-488-856A-3
31	95	6.2	10095	3	US-08-822-586-45
32	94.5	6.2	1227	4	US-09-252-991A-15822
33	94.5	6.2	1353	4	US-09-252-991A-15870
34	94.5	6.2	1719	4	US-09-252-991A-15897
35	94	6.2	1257	4	US-09-489-039A-6981
36	94	6.2	3765	4	US-09-266-965-9
37	94	6.2	53500	4	US-09-266-965-76
38	93.5	6.1	1155	4	US-09-252-991A-14787
39	93.5	6.1	2082	4	US-09-252-991A-2596
40	93.5	6.1	2127	4	US-09-818-780-75
41	93.5	6.1	2388	4	US-09-252-991A-2514
42	93	6.1	2401	3	US-09-320-878-20
43	93	6.1	2401	4	US-09-141-908-9
44	93	6.1	2401	4	US-09-657-440-20
45	93	6.1	3147	2	US-08-781-802-7

#### ALIGNMENTS

#### RESULT 1

US-09-023-655-809  
; Sequence 809, Application US/09023655  
; Patent No. 6607819  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:



TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 809:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1341 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRAITUT01  
CLONE: 753522  
US-09-023-655-809

Alignment Scores:  
Pred. No.: 4,75e-12 Length: 1341  
Score: 193.00 Matches: 88  
Percent Similarity: 40.36% Conservatives: 48  
Best Local Similarity: 26.11% Mismatches: 110  
Query Match: 12.69% Indels: 94  
DB: 4 Gaps: 10

US-10-042-894A-8 (1-289) x US-09-023-655-809 (1-1341)

QY 8 GluHisGlnValAlaGlyHisArgAlaSerAlaSerLys-LeuGlyProLeuIleAspG1 27  
Db 176 AAACATCATGTCTTAGGAAGACAGAAAGAACCCCTAAGGAGCTGGGTGGCTCACAGATAA 235  
QY 27 ySerGlyLeuPheTyLysPro----- 34  
Db 236 AAGAGAGAAATAGAGCCCATAGTAGAAGACGAATTTGAGTGGCTAAAGAAATCTG 295  
QY 35 -----LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPh 51  
Db 296 AAGTCTTGTACTACACTGTAGAGAGAGAGGGAATATAAG-----TT 337  
QY 51 eSerAla-----HisAlaAlaValProAlaArgIleArgAspTh 64  
Db 338 CCAGCTTAACACTATAACCTTGGAGCATGAATGTCCAGCAGCAACAGTTACAGAGAA 397  
QY 64 rPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGluPr 84  
Db 398 T-----GAGGAGATGCAAGCATCGGAAC 424  
QY 84 o-HisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysValAlaA 104  
Db 425 AGTACAAATTTATCTTACTGGAAACCTGACTTCCCGCTATGAGTGCCTTGTGCTCTTG 484  
QY 104 spTlelylledlyAlalleThrTrpProSerProGluProTyrlleAlaLyBC 124  
Db 485 ACCTCAAGTGGGCAACAGCAACATGGTGTATGATGCTTCAGAGGAGAGAGCCCAACC 544  
QY 124 ysLeuAlaLysAspArgGlyThrSerValLeuLeuGlyPheArgValSerGlyValA 144  
Db 545 AGATCCGAAATCTCAGCAGAGACATCTCAGTCATTTGGTGTGCTGTGGCATGC 604  
QY 144 rgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluValValysAlaMetA 164  
Db 605 AGGTGTAC-----CAAGCAGGCGAG-TGGGAGCTCATGTTTCATGAACA-AGTACCATTG 656  
QY 164 spThrAlaGlyValArgArgValLeuArgArgTyValSerSerValAlaAspGluGlyM 184  
Db 657 ACGAAGCTATCGGTGCGAGGCTTCAAGAGGACACTTTTCCAGTTCTTCCACATGGGCG 716  
QY 184 etAspCysAlaLeuAlaAlaValTyrglyGlyLysGlyValLeuSerGlnLeuA 204  
Db 717 GTACTCGCGCGTG-----AACTCCTG-GGCCCTGTGCTCAAGAAGCTGA 760  
QY 204 rgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTySerAlaSerIleL 224  
Db 761 CTGAGCTCAAGGCGAGTGTGGAGCGACAGGAGTCTTACCCTTCTACTCAAGTCCCTGC 820  
QY 224 euLeuGlyTyAsp----- 228

Db 821 TGGTCATTATATGTCGCAAGGAGCGCCGAGTGGTCTCTGAGCTCAGATGCTGAGGATT 880  
QY 229 -----AlaAlaValAlaAlaGlyGly----- 236  
Db 881 TGGAGGACCTGTCTCAGAGGAATCAGCTGATGAGTCTGCTGCTGCTATGCTTACAAACCA 940  
QY 237 --AspGlyGlyGlyValThrValLysLeuValAspPheAlaHis----- 250  
Db 941 TCGGCGCCAGCTCTGTAGATGTGCGCATGATGCGATTGTCGACTTTGGCACACACCCTGCAGGCTGT 1000  
QY 251 -----ValAlaGluGlyAspGlyValIleAspHisAsnPheLeuGlyG 265  
Db 1001 ATGGCGAGGACACCGTGTGTCATGAGGGCCAG-----GATGCTGCTATATCTTCG 1051  
QY 265 lyLeuOysSerLeuIleLysPheValSerAepIleValProGluThr 280  
Db 1052 GGCTCCAGAGCGCTGATAGACATTGTTCACAGAGATAAGTGAAGAGAGT 1098

RESULT 2  
US-09-620-312D-712  
; Sequence 712, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Fonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yungding  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 712  
; LENGTH: 1748  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (187)..(1467)  
US-09-620-312D-712

Alignment Scores:  
Pred. No.: 1.99e-11 Length: 1748  
Score: 189.00 Matches: 71  
Percent Similarity: 45.27% Conservatives: 39  
Best Local Similarity: 29.22% Mismatches: 83  
Query Match: 12.43% Indels: 53  
DB: 4 Gaps: 6

US-10-042-894A-8 (1-289) x US-09-620-312D-712 (1-1748)

QY 78 GluAlaGlnProGlyGluPro-HisProHisLeuValLeuAspLeuAlaGlyPh 97  
Db 765 GAATGCAAGCATCGGAGACAGTACAAATTTATCTTACTGGAAACCTGACTTCCGCTA 824

QY	97	eGlnAlaProCysValAlaAlaAspIleValIleGlyAlaIleThrTipProProSerSerPr	117
Db	825	TKAGGTGCCCTTGTCCTTCACTCCAGATGGGCACACGACCAACATGGTCATGATGCTTC	884
QY	117	oGluProTyrlleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuG	137
Db	885	AGAGGAAGGACGCCAACCATGCGAAAAATGTCAGCAGACACATCTGCAGTCATTGG	944
QY	137	yPheArgValSerGlyValargValValGlyProGluGlyAlaValTrrpargThrGluAr	157
Db	945	TGTGCGTGTGTGTGGCATGACAGGTGTAC-----CAAGCAGGCAG-TEGGCAGCTCATGT	997
QY	157	gProGluValLysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValse	177
Db	998	TCATGAACA-AGTACCATGACGAGCTATCGGTGCAGGCTTCAAGGAGGCATTTTC	1056
QY	177	rSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysG	197
Db	1057	CAGTTCTCCACAATGGCGGTACTCGCGCGT-----AACTCCTG-GG	1100
QY	197	yGlyValLeuSerGlnLeuArgGluLeuLysAlaTrrpPheGluGluGlnThrLeuPheHi	217
Db	1101	CCCTGTGCTCAAGAGCTGACTGAGCTCAAGGCAGTGTGTGAGCGACAGAGTCTTACC	1160
QY	217	sPheTySerAlaSerIleLeuLeuGlyTyrAsp-----	228
Db	1161	CTTCTACTCAGCTCCCTGCTGGTCTATTATGATGCCAAGAGCGGCCGAAGTGGTCT	1220
QY	229	-----	235
Db	1221	GGACTCAGATGCTGAGGATTTGGAGACCTGTCAGAGAAATCAGCTGATGAGTCTGCTGG	1280
QY	235	yGly-----AspGlyGlyValThrValLysLeuValAspPheAl	249
Db	1281	TGCCTATGCTTACAAACCCATCGCGCCAGCTCTGTAGATGTGGCATGATCGACTTTC	1340
QY	249	aHis-----ValAlaGluGlyAspGlyValIl	258
Db	1341	ACACACCACCTGCAGCTGTATGGCAGGACACCGTGTGTCATGAGGGCCAG-----	1392
QY	258	eAspHisAsnPheLeuGlyGlyLeuCysSerIleuIleLysPheValSerAspIleValPr	278
Db	1393	-GATGCTGGCTATATCTTCGGGCTCCAGAGCCTGATGACATTTGTACAGAGATAAGTGA	1451
QY	278	oGluThr	280
Db	1452	GGAGAGT	1458
RESULT 3			
US-09-252-991A-6525/G			
; Sequence 6525, Application US/09252991A			
; Patent No. 6551795			
; GENERAL INFORMATION:			
; APPLICANT: Marc J. Rubenfield et al.			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.136			
; CURRENT APPLICATION NUMBER: US/09/252,991A			
; CURRENT FILING DATE: 1999-02-18			
; PRIOR APPLICATION NUMBER: US 60/074,788			
; PRIOR FILING DATE: 1998-02-18			
; PRIOR APPLICATION NUMBER: US 60/094,190			
; PRIOR FILING DATE: 1998-07-27			
; NUMBER OF SEQ ID NOS: 33142			
; SEQ ID NO 6525			
; LENGTH: 1899			
; TYPE: DNA			
; ORGANISM: Pseudomonas aeruginosa			
US-09-252-991A-6525			
Alignment Scores:			
Pred. No.:	0.00696	Length:	1899

Score:	113.50	Matches:	93
Percent Similarity:	33.51%	Conservative:	35
Best Local Similarity:	24.35%	Mismatches:	115
Query Match:	7.46%	Indels:	139
DB:	4	Gaps:	20

US-10-042-894A-8 (1-289) x US-09-252-991A-6525 (1-1899)

QY	12	AlaGlyHisArgAlaSerAlaSerLysLeuGlyProLeuIleaspGlySerGlyLeuPhe	31
DB	1515	GC CGCGCACCGCTACTGTGTGGCTGTAGCGCT-----CATGCGGGCTGGC	1465
QY	32	TyrLysProLeuGlnAlaGlyAsp-----ArgGlyGluHisGluVal	45
DB	1464	GAACAGCCCGAAGACCGCTGGAGACATGTACACCGTGGCCGCTGCGGAGCGCTTCAGGTG	1405
QY	46	AlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArgAspThrPhe	65
DB	1404	CGAGAGAACGCTTGTGTGGCTTACACGCGCCAGAAAGTTGATGTTTCATGATCCACTC	1345
QY	66	PheProArg-----PheHisGlyThrArgLeuLeuProThrGluAlaGlnPro	81
DB	1344	GTAATCCGAGTAGTGCCTTCCACGGTACCGCGTGGCGGCGACCCGCGTTGTGAA	1285
QY	82	GlyGluProHisPro-----HisLeuVal	89
DB	1284	AATCAGGTTACCCGGCCATGTTGCTGGCCGCTTGTGCGCCAGCGCTGCACCTGCTC	1225
QY	90	-----LeuAspAsp-----LeuLeu	94
DB	1224	GCGTCCGACGTCGACGACCGCGGTGGAGACGCTACCCGCTGCTGGCGAGCAATTG	1165
QY	95	AlaGlyPheGlnAlaProCysValAlaAspIleYsIleGlyAlaIleThrTrpProPro	114
DB	1164	GCGGTTTCTCGAGAGTGGCGATTGAGTGGCGGAGGCGCCAGGTGACACCCCTGCT	1105
QY	115	Ser--SerProGluProTyrIleAlaLysCysLeuAlaLysAspArg-----	129
DB	1104	GCCAGCTCCACGGCGGGCGGACCGATCCGCGGCGCTGCGCGGAGTGGCGGGCGAC	1045
QY	130	-----GlyThrThrSer-----ValLeuLeuGlyPheArgValS	141
DB	1044	TTTGTTTCGAACGACTTCARGCGGGTACCCCTTCGCGCTGTTGGCAGGCTTCGTTGCTC	985
QY	141	erGly-----ValArgValValGly-----	147
DB	984	AGGCGCGGACGGTGTGCTGTGAGGTGTGTAGCGGAGGCGTGCAGCTGCGGCTCAGG	925
QY	148	-----ProGluGlyAlaValTrpArgThrGluArgProG	159
DB	924	GCACGMAAGCCAGGTGATCCCGGCGACACAGGTCAGTTGCGGCGCGCTACCGGATGC	865
QY	159	luValLysAlaMetAspThr-----AlaGlyValArgArgValLeuArgA	174
DB	864	AGGTACAGAGCTTCTTGCACACCGCCACCGCTCCAGACGCTGCGCTCCAGTTCGGTGGAGG	805
QY	174	rgTyrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAla-----AlaVal	192
DB	804	TATTGTTGTAGGTTCTCGACCGGGGCTTGACCTCGAGACTGAGCAGTTCGCGGCGC	745
QY	192	alTyrGly-----	194
DB	744	TTTCATCATGCTTCAGCGCATCAGGAGCAGTGGCGCATCTCGACTCGATCATGTAGACCATC	685
QY	195	-----GlyLysGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluG	212
DB	684	GAGTTGTGGCGGAGCCGGTGTTCGGTCCCAT-----CAGG	649
QY	212	IuGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAla-----	229
DB	648	ARGAACAGTTGGGAAAGCCGGCGGTGTAGTGGCTTGTAGGCTTCGGGCGCTTGGTGC	589
QY	230	-----AlaAlaValAla-----AlaGlyGlyAsp-----	237

```

Db      588 CAGCTGTCCAGCAGTCCGGCCGCTGCGCGACCGAGACCACTCGCGGGGATAGGATCG 529
      |||||
QY      238 --GlyGly-----GlyValThrValLysLeuValAspPheAlaHisValAla----- 252
      |||||
Db      528 TTGGCGGTGAAGCCGGTACCAGAGATGATCGCGTCGACCTCGCGTTCCTCCGCGCGTTCGCG 469
      |||||
QY      253 -----GluGlyAspGlyValIle-----AspHisAsnPheLeuGlyG 265
      |||||
Db      468 TCGACGATTCCATTGGCGGTGACGCGCGCGGATGCTTCGGTGATCACCCTGGAGTTGGCG 409
      |||||
QY      265 ly 265
      ||
Db      408 GC 407

RESULT 4
US-09-252-991A-12872/c
; Sequence 12872, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12872
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12872

```

```

Alignment Scores:
Pred. No.: 0.0123 Length: 984
Score: 107.50 Matches: 91
Percent Similarity: 35.86% Conservative: 32
Best Local Similarity: 26.53% Mismatches: 117
Query Match: 7.07% Indels: 105
DB: 4 Gaps: 15

US-10-042-894A-8 (1-289) x US-09-252-991A-12872 (1-984)
QY      8 GluHisGlnValAlaGlyHisArgAlaSer-----AlaSerLysLeuGlyPro 23
      |||||
Db      968 GAAACGCAATGGCGGAGCAGCTCTACGTGACCGCGCATGGGAATCTACTGGCGGCGG 909
      |||||
QY      24 LeuIleAspGlySerGlyLeuPheThrLysProLeuGlnAlaGlyAspArgGlyGlu--- 42
      |||||
Db      908 TCCGCGGACGCGGAGCAGTATTT-----ACGGTCAGCGCTCGATTTTCGCGTCCCTGT 855
      |||||
QY      43 -----HisGluValAlaPheThrGluAlaPhe----- 51
      |||||
Db      854 CCAGCGTGTAAAGCAGATTTCAGGAGGAGCAGCGCGCTCTCCGCGCTTCGAGGCTGA 795
      |||||
QY      52 -----SerAlaHisAlaValProAlaArgIleAr 62
      |||||
Db      794 TCCGCGCGTCCGCGGACGCGCTGCTGCTGACGACGCGGAGCAATTCACGCTGTTCTCCG 735
      |||||
QY      62 gAspThr-----PhePheProArgPheHisGlyThrArgLeuLeuProThrG1 78
      |||||
Db      734 GCGTACGCGGCGGCGCAGTACCACATTTCGCGCAGCTGTGTGGA-----ATGACTGGCG 684
      |||||
QY      78 uAlaGlnProGlyGluProHisProHisLeuValLeu----- 90
      |||||
Db      683 AGAACAGCCCGGTCCCGGCGCGGACCTCTGCTGCTCCAGCGCCGAGGATACAGGC 624
      |||||
QY      91 -----AspAspLeuLeuAlaGlyPheGlnAl 99
      |||||

```

```

Db      623 GGCGCTTACGCGCGCGAGATATGCTATGCGGCGGACGATTTCTCTGCGCGTATAGCAGC 564
      |||||
QY      99 aProCysValAlaAsp----- 105
      |||||
Db      563 CCTTCTTGAAGCTGACCCCGCGGACCGCTGCGAGTTGATCATCTGTGGGATGAACAGCT 504
      |||||
QY      105 e-LysIleGlyAlaIleThrProSerSerProGluProTyrlleAlaLysCysL 125
      |||||
Db      503 CGCGGTAGGCGCGAACCACCTGGCG-----ATGCTGCG- GCGACCTGCGCGGAGCAGC 451
      |||||
QY      125 euAlaLysAspAspGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgV 145
      |||||
Db      450 CAATCGTCGAGCGCGCTCGCGCTGTGTCTGCG- GGAGCGTGGCAAGCAGCGCTGGCG 392
      |||||
QY      145 alValGlyProGluGlyAlaValTrpArgThrGluArgPro---GluValLysAla-Met 163
      |||||
Db      391 GCGTTGCGCGGAAACCATAGCTCGACGCGCGCGCTCGCGCTCTCCCGGGGTGCTATATCCCGGCGCGGAGTGC 332
      |||||
QY      164 AspThrAlaGlyValArgArgValLeuArgArgTyVal----- 176
      |||||
Db      331 GCCGTGCGCGCGATCTCGCGCTCTCGCGGGGTGCTATATCCCGGCGCGGAGTGC 272
      |||||
QY      177 -----SerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAla 191
      |||||
Db      271 CTCGCTGCGTGGCGCAAGCTATGCGGAGCCAGGC-----TGGGCTCTCTGCGGCAAG 218
      |||||
QY      192 ValTyrglyGlyGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGlu 211
      |||||
Db      217 GCTGCGCTTGGAGAATACGGGTATTTCTCAGGTCCGCCAATCTGGGCTC----- 167
      |||||
QY      212 GluGlnThrLeuPheHisPheTySerAlaSerIleLeuLeuGlyTyAspAlaAlaAla 231
      |||||
Db      166 GAGCAGCTCGCGCGCATCGCCAGCAGCAGCGCGCTCGCGCTCGCGGCGAGATCGGAGAGCT 107
      |||||
QY      232 -----ValAlaAlaGlyAspGlyGlyGlyValThrValLys 244
      |||||
Db      106 CGATAGCAGCGCGCTTGTATGTCAGCGCGCGGAGGCTGGAGGTCTCGTCTGTGAG 47
      |||||
QY      245 LeuValAspPheAlaHisValAlaGlyAspGlyValIleAspHisAsnPheLeuGly 264
      |||||
Db      46 -----GTAGCGAGGTTCGAGGTTCAGGTGAGTGGCCCTCGGAGCACTGGC 5
      |||||
QY      265 Gly 265
      |||||
Db      4 GGC 2

RESULT 5
US-09-016-434-1132/c
; Sequence 1132, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: 2582 AGTCTCAGCGCCAACTGCAGGCGGTAGCCAGCACCTGCTGCTCTCGCGGTGCAGCGC 2523  
FILING DATE: 186 sAlaLeuAlaAlaValTyrGlyGlyLeuSerGlnLeuArgGluLe 206  
CLASSIFICATION: 2522 GGGGCTCCAGCGTGCAGCAGGCTGCCCTCGGGGCTGTGTGACCGA----- 2474  
ATTORNEY/AGENT INFORMATION: NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION: TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1132:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4078 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g1407624  
US-09-016-434-1132

Alignment Scores:  
Pred. No.: 0.104 Length: 4078  
Score: 107.50 Matches: 81  
Percent Similarity: 33.53% Conservative: 31  
Best Local Similarity: 24.25% Mismatches: 112  
Query Match: 7.07% Indels: 110  
DB: 4 Gaps: 17

US-10-042-894A-8 (1-289) x US-09-016-434-1132 (1-4078)

QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34  
Db 3077 AGTTCTCGCGGAGAGCTTGTCCCGCAGACAGCTCGGTAGCGGA----- 3033

QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54  
Db 3032 ---CGGGGGGGTGAAGGGGGCCCGCGCGGTGTGGGCTCGCACCTTGAGGTGAT 2976

QY 55 -----AlaAlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70  
Db 2975 AGCCGCTGTCGGCTTCAGCCCTGCAGCGTGCAGCGTTCT----- 2936

QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81  
Db 2935 -----CCGCGCGCGCTCAGCGCGCTCGCGCAGCTCAGTCTCTCGGG 2892

QY 82 -----GlyGluProHisProHisLeuValLeuAspAspLeuAlaG1 96  
Db 2891 CAGGGCCAGGCGACCGCGCTCCCGCAGCGCCACCGTGTATTGAGAT----- 2843

QY 96 yPheGlnAlaProCysValAlaAspIleGlyAlaIleThrTrpProProSerSe 116  
Db 2842 -----GGCCCCGTGCGCTCGCGCGGCGCGGTGCAGCAGCGGA 2802

QY 116 rProGluProTrpIleAlaLysCysLeuAlaLysAspArgGlyThrSerValLeuLe 136  
Db 2801 GAAGGAC-----GGTCCCGCGCGAGGGGTTCGCGCGCGCTCCA 2763

QY 136 uGlyPheArgValSerGlyValArgValVal----- 146  
Db 2762 GAATCTCGGGGTGCCACCGGGGTGTCTCCGGGATGCTCAGGACCTCGGCTGCTCCT 2703

QY 147 -----GlyProGluGlyAlaValTrpArgThrGluArgProGluVally 161  
Db 2702 CGCCAGGCGCGCGGCTCGGGCGCCAGCCGGAACATACGTGGCCCTTGTGCA 2643

QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170  
Db 2642 CGCTGATCCGTGTACGGTCTCCGAGGGCGGGAACCTCCAGGTGGCCAGGGGCTCG 2583

QY 170 gValLeuArgArgTyrValSerValAlaAsp-----GluGlyMetAspCy 186  
Db 170 -----GluGlyMetAspCy 186

Db 2582 AGTCTCAGCGCCAACTGCAGGCGGTAGCCAGCACCTGCTGCTCTCGCGGTGCAGCGC 2523  
QY 186 sAlaLeuAlaAlaValTyrGlyGlyLeuSerGlnLeuArgGluLe 206  
Db 2522 GGGGCTCCAGCGTGCAGCAGGCTGCCCTCGGGGCTGTGTGACCGA----- 2474  
QY 206 uLysAlaTrpPheGluGluGln---ThrLeuPheHisPheTyrSerAlaSerIleLeuLe 225  
Db 2473 -CAGGGTGGCGCGCCAGCAGCTCTCTCTGCTGTCACACAC-----CT 2430  
QY 225 uGlyTyrAspAlaAlaAlaValAlaGlyGlyAspGlyGly----- 239  
Db 2429 TGGGTTTTCGCGAGCGCCATCGCTTCATGTGTAGCGGCTACCGTGTAGGTAGTACG 2370  
QY 240 -GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAs 259  
Db 2369 CGGTCTCAGGCTGCAAGTTTGTGAT---GACCATCTCCTG-----GGCATCGGCCA 2322  
QY 259 pHisAsnPheLeu-----GlyGlyLeuCysSerLeuIleLysPheValSe 274  
Db 2321 GCATGACGTCTTGTATGCTGGCGCGCGCGGCTCGCGCTCTCCATGCGCAGCTAGT 2262

QY 274 rAspIleValProGluThrProHisThrGlnProLeuGly 287  
Db 2261 GGACTGTGTAGCGCGGATCTGCGCGTGTGCGCGCGGG 2222

RESULT 6  
US-08-348-006B-6/c  
Sequence 6, Application US/08348006B  
Patent No. 5658756  
GENERAL INFORMATION:  
APPLICANT: RODAN, GIDEON A.  
APPLICANT: SCHMIDT, AZRIEL  
APPLICANT: RUTLEDGE, SU JANE  
TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN  
TITLE OF INVENTION: TYROSINE PHOSPHATASE  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: J. MARK HAND  
STREET: 126 E. LINCOLN AVE., P.O. BOX 2000  
CITY: RAHWAY  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/348,006B  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/122,032  
FILING DATE: 14-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: HAND, J., MARK  
REGISTRATION NUMBER: 36,545  
REFERENCE/DOCKET NUMBER: 189921A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-3905  
TELEFAX: 908-594-4720  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6000 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
US-08-348-006B-6

Alignment Scores:

```
Pred. No.: 0.24 Length: 6000
Score: 106.50 Matches: 80
Percent Similarity: 33.64% Conservative: 31
Best Local Similarity: 24.24% Mismatches: 117
Query Match: 7.00% Indels: 102
DB: 1 Gaps: 15

US-10-042-894A-8 (1-289) x US-08-825A-6 (1-6000)

QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34
Db 3077 AGTTCTTGGCGAGACTTGGTCCCGCAGGAACGTCCGGTAGCGGA----- 3033
QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54
Db 3032 ---CGGGGGGCTGAAGGGGCCAGGGCCCCGGCCCGTGTGGGCTCCGACTTGGAGGTCTAT 2976
QY 55 -----AlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70
Db 2975 AGCCCGTGTGGGCTTTCAGGCCCTTCAGCGCTGAGCGTCTCT----- 2936
QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81
Db 2935 -----CGGCGCCGGCTCAGCGCGCTCGCGCAGCTCAGTCTCTCGG 2892
QY 82 -----GlyGlnProHisProHisLeuValLeuAspLeuLeuAlaG 96
Db 2891 CAGGGCCAGGCGACCGCGCTCCCGCAGCGCACCGTGTATTGACGAT----- 2843
QY 96 yPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProSerse 116
Db 2842 -----GGCCCCGTGTGCGTCCGGCGGCACGGGTGGGACGACCGGA 2802
QY 116 rProGluProTyIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLe 136
Db 2801 GAAGGAC-----GGTCCCGCGGAGGGGTTCGCGCGCGCTCCA 2763
QY 136 uGlyPheArgValSerGlyValArgValVal----- 146
Db 2762 GAATCTGCGGGTGGCCACGCGGGGTGTCTCCGGGATGCTCAGGACCTCGGCTCGCTCCT 2703
QY 147 -----GlyProGluGlyAlaValTrpArgThrGluArgProGluVally 161
Db 2702 CGCCAGGCGCGGGCTCCGGCGCGCAAGCGGAACATAGTGGCCCTTGTGCA 2643
QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170
Db 2642 CGCTGTATGCGGTAGCGGTCTCCGAGGGCGGGAACCTCCAGGGTGGCCAGGGCGGTGCG 2583
QY 170 gValLeuArgArgTyrValSerSerValAlaAsp-----GluGlyMetAspCy 186
Db 2582 AGTCCTCACGGCCAAACTGTCAGGCGGTAGCCAGCACCTGGTCTCCGCGGTGCCAGCGG 2523
QY 186 sAlaLeuAlaAlaValTyrGlyGlyLysGlyValLeuSerGlnLeuArgGluLe 206
Db 2522 GGGGCTCCAGGTGTCAGCAGCGCTGCCCTCGGGGTCTCTGTCACCGA----- 2474
QY 206 uLysAlaTrpPheGluGluGln---ThrLeuPheHisPheTyrSerAlaSerIleLeuLe 225
Db 2473 -CAGGGTGGCGGCCCCAGCAGCTGCTCCCTGGTGTCAACACAC-----CT 2430
QY 225 uGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly----- 239
Db 2429 TGGGTGTGCTCGAGCGCCATCGCCCTTCATGTGTGTAGCGGGTACCGTATGAGGTACG 2370
QY 240 -GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyValAsp---GlyValI 258
Db 2369 CGGTCTCAGGTCTGAAGTTGTGATGACCATCTCTCGGCATTCGGCAGCATGACGTCT 2310
QY 258 eAspHisAsnPheLeuGlyLeuCysSerLeuIleLysPheValSerAspIleValPr 278
Db 2309 TGATCGCGGGCGCGCGGCGCTCGCGCGCTCCATGCGGACGTAGTGGACCTGGTAGC 2250
```

```
QY 278 oGluThrProHisThrGlnProLeuGly 287
Db 2249 CGCGGATCGCGCGTGTGCTCGCGCGCGG 2222

RESULT 7
US-08-800-825A-6/c
; Sequence 6, Application US/08900825A
; Patent No. 5866397
; GENERAL INFORMATION:
; APPLICANT: RODAN, GIDEON A.
; APPLICANT: SCHMIDT, AZRIEL
; APPLICANT: RUTLEDGE, SU JANE
; TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSER: J. MARK HAND - MERCK & CO., INC.
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,825A
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HAND, J. MARK
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 18992DA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-3905
; TELEFAX: 732-594-4720
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6000 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-800-825A-6

Alignment Scores:
Pred. No.: 0.24 Length: 6000
Score: 106.50 Matches: 80
Percent Similarity: 33.64% Conservative: 31
Best Local Similarity: 24.24% Mismatches: 117
Query Match: 7.00% Indels: 102
DB: 2 Gaps: 15

US-10-042-894A-8 (1-289) x US-08-800-825A-6 (1-6000)

QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34
Db 3077 AGTTCTTGGCGAGACTTGGTCCCGCAGGAACGTCCGGTAGCGGA----- 3033
QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54
Db 3032 ---CGGGGGGCTGAAGGGGCCAGGGCCCCGGCCCGTGTGGGCTCCGACTTGGAGGTCTAT 2976
QY 55 -----AlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70
Db 2975 AGCCCGTGTGGGCTTTCAGGCCCTTCAGCGCTGAGCGTCTCT----- 2936
QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81
Db 2935 -----CGGCGCCGGCTCAGCGCGCTCGCGCAGCTCAGTCTCTCGG 2892
```



Db 2642 CCGCTGATCCCGTGTAGCGGTCTCCGAGAGCGGGAACCTCCAGGGTGCACCGGGCGCTCG 2583  
QY 170 gValLeuArgArgTyrValSerSerValAlaAasp-----GluGlyMetAspCy 186  
Db 2582 AGTCTCTACGGCCAACTGCAGCGGTAGCCAGCACCTGCTCTCCGCGGTGCCAGCGC 2523  
QY 186 salAlaLeuAlaAlaValTyrGlyGlyslgGlyValLeuSerGlnLeuArgGluLe 206  
Db 2522 GGGGTCCAGCGTCCAGCGGCTCCGAGCGGTCCCTCGGGGTCTGCTGCACCGA----- 2474  
QY 206 uLysAlaTrpPheGluGluGln---ThrLeuPheHisPheTyrSerAlaSerIleLeuLe 225  
Db 2473 -CAGGGTGGCGCGCCACACTGCTCCCTTGGTCAACACAC-----CT 2430  
QY 225 uGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly----- 239  
Db 2429 TGGGTGTGTCGAGCGCCATCGCCCTTCATGCTGTAGCGGCTACCGGTGAGTAGC 2370  
QY 240 -GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp---GlyValI 258  
Db 2369 CCGTCTCAGGCTGCAAGTTTGTGATGACCATCTCCTGGGCATCGGCAGCATGACGCTCT 2310  
QY 258 eAspHisAsnPheLeuGlyGlyLeuCysserLeuIleLysPheValSerAspIleValPr 278  
Db 2309 TGATCGCGGCGCGCGCGGCTCGCGCCCTCCATCGGCACGTAGTGGACCTGCTAGC 2250  
QY 278 oGluThrProHisThrGlnProLeuGly 287  
Db 2249 CCGGATCTGGCGGTCTGCTCGCGCGCGG 2222

RESULT 9  
PCT-US94-10166-6/c  
; Sequence 6, Application PC/TUS9410166  
; GENERAL INFORMATION:  
; APPLICANT: RODAN, GIDEON A  
; APPLICANT: SCHMIDT, AZRIEL  
; APPLICANT: RUTLEDGE, SU JANE  
; TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN  
; TITLE OF INVENTION: TYROSINE PHOSPHATASE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: JOHN W. WALLEN III  
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.  
; CITY: RAYWAY  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/10166  
; FILING DATE: 09-SEPT-1994  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/122,032  
; FILING DATE: 14-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WALLEN, JOHN W III  
; REGISTRATION NUMBER: 35403  
; REFERENCE/DOCKET NUMBER: 18992  
; TELEPHONE: 908-594-3905  
; TELEFAX: 908-594-4720  
; TELEX: 138825  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6000 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

i MOLECULE TYPE: cDNA  
PCT-US94-10166-6  
Alignment Scores:  
Pred. No.: 0.24 Length: 6000  
Score: 106.50 Matches: 80  
Percent Similarity: 33.64% Conservative: 31  
Best Local Similarity: 24.24% Mismatches: 117  
Query Match: 7.00% Indels: 102  
DB: 5 Gaps: 15  
US-10-042-894A-8 (1-289) x PCT-US94-10166-6 (1-6000)  
QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34  
Db 3077 AGTCTTCGGGCGAGACTTGGTCCCGCAGGAACGTCGGGTAGCGGA----- 3033  
QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54  
Db 3032 ---CGGGGGGCTGAAGGGGCCAGGGCCCGCGCGCTGTGGCTCGCACTTGGAGGTCT 2976  
QY 55 -----AlaAlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70  
Db 2975 AGGCGGTCTCGGGCTTCAGGCCCTGCACGCTGAGCGTTCT----- 2936  
QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81  
Db 2935 -----CGCGCCCGGCTCAGCGCGCTCGCGGCGCTCAGTCTCTCGGG 2892  
QY 82 -----GlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGl 96  
Db 2891 CAGGGCCCGAGGACCGGCTCCCGCACCGGCACCGTGTATTGACGAT----- 2843  
QY 96 yPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSe 116  
Db 2842 -----GGCCCGTTCGCTCGCGGGGCGACGGGTGGCAGCCAGCGGA 2802  
QY 116 rProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLe 136  
Db 2801 GNAGGAC-----GGTCCCGCGGAGCGGTTCGCGGCGCTCCA 2763  
QY 136 uGlyPheArgValSerGlyValArgVal----- 146  
Db 2762 GAATCTCGGGGTGGCGACGCGGCGTGTCTCCGGGATGCTCAGGACCTCGGTGCTCTCT 2703  
QY 147 -----GlyProGluGlyAlaValTrpArgThrGluArgProGluVally 161  
Db 2702 CGCCCGAGCGCGCGGCTCCGGGCGGCGAGCGGACACATACGTGGGCCCCCTTGTGCA 2643  
QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170  
Db 2642 CGCCTGATGCCGTGTAGCGGTCTCTCGAGGGCGGGAACCTCCAGGGGTGGCGAGGGCGCTCG 2583  
QY 170 gValLeuArgArgTyrValSerSerValAlaAasp-----GluGlyMetAspCy 186  
Db 2582 AGTCTCTACGGCCAACTGCAGCGGTAGCCAGCACCTGCTCTCCGCGGTGCCAGCGC 2523  
QY 186 salAlaLeuAlaAlaValTyrGlyGlyLysGlyValLeuSerGlnLeuArgGluLe 206  
Db 2522 GGGGTCCAGCGTCCAGCGGCTGCCAGCGGTGCCCTCGGGGTCTGCTGCACCGA----- 2474  
QY 206 uLysAlaTrpPheGluGluGln---ThrLeuPheHisPheTyrSerAlaSerIleLeuLe 225  
Db 2473 -CAGGGTGGCGCGCCACACTGCTCCCTTGGTCAACACAC-----CT 2430  
QY 225 uGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly----- 239  
Db 2429 TGGGTGTGTCGAGCGCCATCGCCCTTCATGCTGTAGCGGCTACCGGTGAGTAGC 2370  
QY 240 -GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp---GlyValI 258  
Db 2369 CCGTCTCAGGCTGCAAGTTTGTGATGACCATCTCCTGGGCATCGGCAGCATGACGCTCT 2310





```
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-132

Alignment Scores:
Pred. No.: 4.65 Length: 36412
Score: 105.50 Matches: 74
Percent Similarity: 35.33% Conservative: 32
Best Local Similarity: 24.67% Mismatches: 119
Query Match: 6.94% Indels: 75
DB: 4 Gaps: 14

US-10-042-894A-8 (1-289) x US-08-311-731A-132 (1-36412)
QY 6 ProProGluHisGlnValalaGlyHisArgAlaSerAlaSerLysLeuGlyPro-----23
DB 16731 CCGATGAATACGCCATTGCACCTAGAAACATCAAGCAAGCGCCGAAAGGCCCTGCTGCA 16672
QY 24 -----LeuIleAspGlySerGly---LeuPhe 31
DB 16671 TGAACACCTTGACGGTGGTGGCCCGCCAGCGCGTGTGGATATTCCCGCCAGGTCGGC 16612
QY 32 TyrLysProLeuGlnAlaGlyAspArgGlyGluHisGluValalaPheTyrGluAlaPhe 51
DB 16611 TATGATAGGCTCCCGCTACCGACGTTGAGTCGCTAGACACATGGTTTCGCACTGCT--- 16555
QY 52 SerAlaHisAlaValProAlaArgIleArgAspThrPheProArgPheHisGly 71
DB 16554 ---TCGACAGCGCGCTGTTGGAGCGCTACCTG-----16525
QY 72 ThrArgLeuLeuProThrGluAlaGlnProGlyGluProHisProHisLeuValLeuAsp 91
DB 16524 -----GAGCGGTTTCTCATACCGTG-----16504
QY 92 AspLeuLeuAlaGlyPheGlnAlaPro-----CysVal 102
DB 16503 -----GCCGTGATGCAGACACACAGAGCGGTTGCACCGCTCGCTTATGAGTCGGTG 16453
QY 103 AlaAspIleLysIleGlyAlaIleThrTrpProPro-----SerSerProGluProTyr 120
DB 16452 GAAGACCTCCCGCTGATTGGTGGTCTAGCCGCAAGTCGGGTTCGCTCGGAGCTGCAC 16393
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 16392 ATCGACGAGGGTTCGTTGATGAGGTCCTGCAAGCGTGTGGGGGTTT---GCT 16336
QY 141 SerGlyValArgValValGlyProGluGly---AlaValTrpArgThrGluArgProGlu 159
DB 16335 GACGGTGAAAGGGCTTGTGCGCGCAGAGGCAATGCCATC-----ACAGTGCCTGCTTG 16282
QY 160 ValLysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerVal 179
DB 16281 GTACAGCGATGCGGACGCGGCAATGTCGCGGAGATCCCGAGTAGCATCCGGTTC 16222
QY 180 AlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGly---198
DB 16221 CGGACAAAGAGTCTC-----GTCGGATTCCACATCGCTGCTGTGAGGCTGGGCAC 16171
QY 199 -----ValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGlu 211
DB 16170 CCTCCGACTCGGCACCTCGACGCTTTTGTAGTACATGCTAGTAAATATGCGCGCTTC--- 16114
QY 212 GluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaA 231
DB 16113 -----ACTATTCATCGGGGTGAGCGGTTCGGCTGCCATCTATCCAGGCGCATC 16063
QY 232 ValAlaAlaGlyGlyAspGlyGlyGlyValThrValLysLeuValAspPheAlaHisVal 251
DB 16062 GCGTTCGCGAGCTGACCGGCTGGGCCACGCGGTGGGATCGTCGACGACATCGAGTT 16003
QY 252 AlaGluGlyAspGlyValIleAspHisAsnPheLeuLeuGlyGlyLeuCysSerLeuIleLys 271
DB -----ATTACGATGCAGACCCAGAGCTGTACGAGAG 1265

Db 16002 GATCCTGGTGGTGGCATC-----CGGTTGGGTCCCTGGCTCGATCCTGGCG 15955

RESULT 12
US-08-637-899-2
; Sequence 2, Application US/08637899
; Patent No. 5908772
; GENERAL INFORMATION:
; APPLICANT: Mitta, Masanori
; APPLICANT: Sano, Mutsumi
; APPLICANT: Kato, Ikunoshin
; TITLE OF INVENTION: Gene Encoding Lacto-N-Biosidase
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,899
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1917 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Streptomyces sp.
; STRAIN: 142
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1917
; IDENTIFICATION METHOD: E
US-08-637-899-2

Alignment Scores:
Pred. No.: 0.0826 Length: 1917
Score: 104.00 Matches: 66
Percent Similarity: 34.73% Conservative: 25
Best Local Similarity: 25.19% Mismatches: 97
Query Match: 6.84% Indels: 74
DB: 2 Gaps: 13

US-10-042-894A-8 (1-289) x US-08-637-899-2 (1-1917)
QY 4 LeuHisProProGluHisGlnValalaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
DB 1149 CTCATCCAGCAGCC---GTCCTCGCTGCTCGCCGAGGCGCGCC-----1190
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
DB 1191 -----GGTCATGAATCCGCCCTCTACCTCTCTGCTGGTGGCGCGCG 1232
QY 44 GluValalaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
DB 1233 -----ATTACGATGCAGACCCAGAGCTGTACGAGAG 1265
```

```
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 1266 -----CGACTGACGCCCTTGGCTTCAGGGGCGAGCGCTGACCCAGGGGCG 1313
QY 84 ProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103
Db 1314 GCGGAACCTCACCGCGCGAGATCAGCTGTGGCGGA----- 1352
QY 104 AspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAlaLys 123
Db 1353 -----CAG 1355
QY 124 CysLeuAlaLysAsp-----ArgGlyThrThrSerValLeuLeuGlyPheArgValSer 141
Db 1356 TCGGGCGCGAGACGGAGACAGGTTCGAGACGAGGTCTTCATGCGCTCGCTTCGT 1415
QY 142 GlyValArgValValGly-----ProGlyValAlaValTrpArgThrGlu 156
Db 1416 GCGCGAGCGACCTGGGGGCGCCGAGCGCGAGCCGACGTACGCGGTTTCGAGCGCT 1475
QY 157 ArgProGluValLysAlaMetAspThrAlaGlyValArgArg-----ValLeuArg 173
Db 1476 CCGCCGGAAGATCGGTACGCGCGCGGCTGGGAGAACACCGACCGCGCGCTCGCCGA 1535
QY 174 ArgTyrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyr 193
Db 1536 CGGTACGTACCG-----GCTGACACGCGGCGGAA-----GGCGCTGGCCCCCACC 1580
QY 194 GlyGlyLysGlyGlyValLeuSerGlnLeuArgGlu-----LeuLysAlaTrpPheGlu 211
Db 1581 GCGCGACGCGCGCTGTCCCTGTCTGTAAGAACAGCGCGGCTCTGGGCGCTGACGCGAC 1640
QY 212 GluGlnThrLeuPheHis-----PheTyrSerAlaSerIleLeuLeuGly----- 226
Db 1641 CCGCGACGCGGTACTACACGCTCGCTCCACGAGAGCGGTCACTGCGTGGCGCGGTGCG 1700
QY 227 ---TyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGlyValThrValLysLeu 245
Db 1701 CGGCAAGAGTACCTGGGCGCGCGCTGAGGTTGGGGGGGAGCTGTCACTCGCAACTG 1760
QY 246 Valasp 247
Db 1761 CTCGAC 1766

RESULT 13
US-09-252-991A-11331
; Sequence 11331, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11331
; LENGTH: 1896
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11331

Alignment Scores:
Pred. No.: 0.0925 Length: 1896
Score: 103.50 Matches: 105
Percent Similarity: 34.54% Conservative: 29
Best Local Similarity: 27.06% Mismatches: 112
Query Match: 6.80% Indels: 145
```

```
DB: 4 Gaps: 23
US-10-042-894A-8 (1-289) x US-09-252-991A-11331 (1-1896)
QY 16 AlaSerAlaSerLysLysLeuGlyProLeuIleAspGlySerGlyLeuPheTyrLys-Pro-- 34
Db 716 TCCGCGCGCCAGCAAACTG-----GCGAGCGGTTGGCGACGAAAGTCCGGG 760
QY 35 -LeuGlnAlaGly--AspArgGlyGluHisGluVal----- 45
Db 761 TCGCGGACAGCGCTCGATCGCGGTAGGCACGGCGCAGCCTCCAGCACGTAATGGGCG 820
QY 46 -AlaPheTyrGluAlaPheSerAlaHisAlaAlaValPro----- 58
Db 821 CGCTGTACCGCTCGGCTTGTGCCACGGCAGCTTTCAGCATCTCGAGGCTCTCGCGC 880
QY 59 -----AlaArgIleArg-AspThrPhe-----PheProArgP 69
Db 881 AGGCGCAGCGCGCGCGCGCGGTGGCTGATCAGTTCGCGGCGGTGGCCACG 940
QY 69 heHisGlyThrArgLeuLeu-----ProThrGluAlaGlnProGlyGluProHisP 86
Db 941 TTCACCTCAGCGCGCGGTTCGATCACCGGTAGTTCGCGCAGGTTCGCGCGC----- 993
QY 86 roHisLeuValLeuAspAspLeuLeuAlaGlyPheGln-----AlaProC 101
Db 994 -----GCTCCAGATGCCGCGCGCGCGCGGACCCCGGTGAGCAGCTTGTCCGCT 1042
QY 101 ysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrI 121
Db 1043 GTCTCGCTTCGTAGAA-----ACCGATCCGCGCGGTAGCGCGGAGGCGTTC 1090
QY 121 leAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal- 140
Db 1091 AGGTACGTGCCA-GCTCGGTTGCGGAGCGAGGTTCATTCTCGTGGGGAATTCGCCCC 1149
QY 141 -----SerGlyValArgValValGlyPro-----GluGlyAlaValTrpA 154
Db 1150 GTCGAGGAAGATCCGCGCGGTTTCGGGTCTTCGCGCATCGCTTCAGCGCA---TGCG 1206
QY 154 rgThrGluArgProGluValLysAlaMetAspThrAlaGlyValArgArgValLeuArgA 174
Db 1207 GCGCGCGTCACGCTAG-ATCCGCTCGACGAGTACCCATCCAC-CGCCAGCGGATCGCG 1264
QY 174 rgTyr-----ValSerSerValAla----- 180
Db 1285 GGCACACAGATTGTCGGCCAGCGGGCTTCGTCGTHAGCGGCTCTCAATTCGACCGCGC 1324
QY 181 -----AspGluGlyMetAspC 186
Db 1325 GCGCGCAGCGCGGGGATCCGCGCGCGCGCGCGGTTCGACGAGCAGCTGGGGTCCGACC 1384
QY 186 ysAlaLeuAlaAlaAlaValTyrGlyGlyLysGlyVal----- 199
Db 1385 TTGCCATT---GCGCGGTACATGTCGCGGTAGGCGGCTTCGCGCGCGGCTCGCGGGG 1441
QY 200 -----LeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuP 216
Db 1442 TCGAGGACGATAGGTGGCTGGCGCGCGCTCGCGCAGAGGAAGAATCCACCGCG 1501
QY 216 heHisPheTyrSerAlaSerIleLeuLeuGlyTyr----- 227
Db 1502 CCGAG-----TCCGAGCGGTAGGCTCGGCTACCGCGAGGCGAGCGCGATGCCCGC 1555
QY 228 -----AspAlaAlaAlaValAlaAlaGly----- 235
Db 1556 GCAGCGTCGAGGCGTTGCCGCAATTGGCCAGGGTTTCAGGCCCGCGCGCGTTCGCGGCG 1615
QY 236 -----GlyAspGlyGly-----GlyValThrValLysLeuValA 247
Db 1616 GGATGTGGCGTGGCGACGCGCGGCTGCTCGGTGCTGGCGGTTGCCGGAAGCTGTGCG 1675
QY 247 spPheAlaHisValAlaGlu-----GlyAspGlyV 257
```

Db 1676 GCATGCGCCGCGAGGTGAGCAGGAGGACAGGCCGCGCGCGACACGAGGGGTACGAGCGGC 1735  
QY 257 allleAspHisAenPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleV 277  
Db 1736 AGTTTCAGGAGACATGCGCGCTCACTGTTTCAGCGCGGTTCGCGGACGGG----- 1788  
QY 277 alProGluThrProHisThr 283  
Db 1789 --CGCGAAGCGCTCAGACT 1806  
RESULT 14  
US-09-252-991A-12565/c  
; Sequence 12565, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 12565  
; LENGTH: 960  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-12565  
Alignment Scores:  
Pred. No.: 0.0724 Length: 960  
Score: 100.50 Matches: 68  
Percent Similarity: 37.60% Conservative: 23  
Best Local Similarity: 28.10% Mismatches: 83  
Query Match: 6.61% Indels: 70  
DB: 4 Gaps: 9  
US-10-042-894A-8 (1-289) x US-09-252-991A-12565 (1-960)  
QY 80 GlnProGlyLeuProHisProHisLeuValLeu----- 90  
Db 759 CAGCCCGGTCCGCGCGCGCCGACCTCGTCTGCGTCCAGCGCCAGCGATACAGCGCGCG 700  
QY 91 -----AspAspLeuLeuLaglyPheGlnAlaPro 100  
Db 699 CTTGAGCGCGCGAGATGATGATCGCGCGCGCGCTGATCATCTGTGGATGACAGCTGCG 640  
QY 101 CysValAlaAsp-----Ile-Ly 106  
Db 639 CTTGAGCTGACCCCGCGCGCGCGCTGCGCTGATCATCTGTGGATGACAGCTGCG 580  
QY 106 sileGlyAlaIleThrTrpProProSerSerProGluProTyrIleAlaLysCysLeuAl 126  
Db 579 GGTAGCGCGCGAACAACCTGGCG-----ATGCGCTGC-GCGCAGCTGGCGCGAGCGCAAT 527  
QY 126 aLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgValVa 146  
Db 526 CGTCGAGCGCGCGCTCGCGCTGTGTGTCG-GGAGCGGTGGCAAGCACCGCTCGCGCGCGT 468  
QY 146 lGlyProGlyGlyAlaValTrpArgThrGluArgPro--GluValLysAla-MetAspT 165  
Db 467 TGGCGCGGACCCCTAGCTACGCGCGCGCTGCGCCAGCGCGCGAGCGCGAGCGTGC 408  
QY 165 hrAlaGlyValArgArgValLeuArgArgTrpVal----- 176  
Db 407 TGGCGCGCGATCTGCGCGCTCTCGCGCGGGGTGTCTATCCCGAGCGCGCGAGTGCCTCG 348  
QY 177 -----SerSerValAlaAspGlyMetAspCysAlaLeuAlaAlaValT 193

Qy	50	AlaPheSerAlaHisAlaAlaValProAlaArgIleArgAspThrPhePheProArg---	68
Db	1169	TCCCGCGGACGAGTGGCTCTGCTCCACGCACTGTCCGACGTGGAGAGGTCCCGCGGGAT	1110
Qy	69	PheHisGly-----ThrArgIleuLeuProThrGluAlaGlnPro	81
Db	1109	GATCATGGGATGGAGTACTGTCAGCTTCTCGGGCTCTTGTACCAAGGCACGAGCACAT	1050
Qy	82	GlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCys	101
Db	1049	GGACTGGAGTGCAGCACCTCCGGTA---GAGTTGGGAAGCCGCGCGCGCTGCG	993
Qy	102	-ValAlaAspIleLysIleGlyAlaIle-----ThrTrpProProSerSerProGluProTy	120
Db	992	CGCGTGGATGAGCGGTGTCCGTAGTGTGGCGCTGCCACCCCTCCGCGCGCGCGCG	933
Qy	120	rIle-----AlaLysCysLeuAlaLysAspArgGlyThrThrSerVa	134
Db	932	CTGCCCGTTGGCGTGTAGCAGCGCGCGGTGCTCGCGC---TCGCGCTTCCTCGCTCGGC	876
Qy	134	LLeuLeuGlyPheArgValSerGlyValArgValValGlyProGluGlyAlaValTrpAr	154
Db	875	GTTTCATGTCATGAAGCGCGACCCACGAGGTTGAG-----GAAGCGCCCATGATCCG	823
Qy	154	gThrGluArgProGluValLysAlaMetAspThrAlaGlyValArgArg-----	170
Db	822	TGAGGCCGCTAAGGATGTAGACGAAGCTGAAGGCCACGTACTCGGCTGGCTGTGCGAGG	763
Qy	171	-----ValIleuArgArgTyValSerSerValAlaAspGluClyMetAspCysAlaLe	188
Db	762	CTTGGTCTCTTCGAGCGCCACGTCGAGTCGCGAAGCCGAT-----	723
Qy	188	uAlaAlaAlaValTyGlyGlyLysGlyGlyValLeuSerGlnLeuArgGluLeuLysAl	208
Db	722	-----GGTGGTGAGGT-----	711
Qy	208	aTrpPheGluGluGlnThrLeuPheHisPheTySerAlaSerIleLeuLeuGlyTyArS	228
Db	710	-----GATGAAGCAGTAGTAGTAGAGCCTGGA	685
Qy	228	pAlaAlaAlaValAlaAlaGlyAspGlyGlyGlyValThrValLysLeuValAspPh	248
Db	684	AGAAGGTCCAGTGTCTGTAGTGGGAGAAGCGCGCGCGCATGCCACAGCGTGTCTATGC	625
Qy	248	eAlaHisValAlaGluGlyAspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSe	268
Db	624	ACGAGAGAAGCCGAT---CAGCACCATGTGTGGCCATGGACACGTCCGC-----	579
Qy	268	rLeuIleLysPheValSerAspIleValProGluThrProHisThrGlnProLeu	286
Db	578	-----GGCGCGCATGCCAGCCGCTT	558

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 27, 2004, 04:36:55 ; Search time 477 Seconds  
(without alignments)  
2573.856 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVAGHRASAK.....IKFVSIVPEPHTQPLGPS 289

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_epool/US10042894/runat\_24032004\_145046\_16937/app\_query.fasta\_1.455  
-DB=N Geneseq 29Jan04 -QMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTENT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10042894 -CGN 1\_1\_470 @runat\_24032004\_145046\_16937 -NCPU=6 -ICPU=3  
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=1 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04: \*  
1: geneseq1980s: \*  
2: geneseq1990s: \*  
3: geneseq2000s: \*  
4: geneseq2001as: \*  
5: geneseq2001bs: \*  
6: geneseq2002s: \*  
7: geneseq2003as: \*  
8: geneseq2003bs: \*  
9: geneseq2003cs: \*  
10: geneseq2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1521	100.0	1344	6	AAD43514
2	1485	97.6	922	6	AAD43513
3	1435	94.3	923	6	AAD43512
4	1406	92.5	1169	6	AAD43511
5	1307	85.9	3416	6	AAD43522
6	1233	81.1	899	6	AAD43518
7	808.5	53.2	643	6	AAD43519
8	742.5	48.8	1020	6	AAD43517

102  
-a"t-  
10 good

9	722.5	47.5	1195	6	AAD43516
10	714.5	47.0	1130	3	AAC48750
11	700.5	46.1	1130	3	AAC40313
12	674.5	44.3	1104	3	AAC33685
13	674.5	44.3	1168	3	AAC39023
14	674.5	44.3	1243	3	AAC38692
15	644.5	42.4	1105	6	AAD43515
16	464	30.5	519	6	AAD43520
17	329	21.6	353	6	AAD43521
18	306.5	20.2	484	9	ADE82058
19	288	18.9	2930	4	ABL23306
20	273.5	18.0	464	6	ABL193575
21	265	17.4	876	4	ABL23307
22	201.5	13.2	2212	7	AAD53176
23	192	12.6	2391	9	ADC06742
24	192	12.6	2890	6	AAD38858
25	191	12.6	2608	7	ADA53453
26	189	12.4	1340	4	AAI60607
27	189	12.4	1737	5	AAH64793
28	189	12.4	1737	9	ADC06746
29	189	12.4	1739	4	AAH18813
30	189	12.4	1748	4	AAI58821
31	189	12.4	1748	8	ADB48802
32	189	12.4	1757	5	AAH64800
33	189	12.4	1990	3	AAH64898
34	183	12.0	1918	3	AAH77195
35	182.5	12.0	1222	9	ADC06748
36	182	12.0	1863	7	ABX70876
37	178.5	11.7	1893	7	ABX70875
38	168	11.0	1723	9	ADC06744
39	168	11.0	4461	6	ABK83730
40	156	10.3	6223	4	ABL05232
41	151	9.9	3197	4	ABL05244
42	145.5	9.6	1782	6	ABL65412
43	145.5	9.6	1782	6	ABL64777
44	145	9.5	1386	6	ABL56444
45	143	9.4	1241	4	AAI87016

#### ALIGNMENTS

RESULT 1  
AAD43514  
ID AAD43514 standard; DNA; 1344 Bp.  
XX  
AC AAD43514;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #4.  
XX  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; ds.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 52..921  
FT /\*tag= a  
FT /product= "Maize IPPK protein #4"  
XX  
FN WO200259324-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 09-JAN-2002; 2002WO-US003120.  
XX  
PR 12-JAN-2001; 2001US-0261465P.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX

```
DR WPI; 2002-636540/68.
DR P-PSDB; AAE26196.
XX
PT New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
XX Claim 1; Page 64-65; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA
XX
SQ Sequence 1344 BP; 228 A; 426 C; 399 G; 291 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      8.13e-131      Length:      1344
Score:          1521.00        Matches:      289
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:     100.00%      Indels:       0
DB:              6           Gaps:       0

US-10-042-894A-8 (1-289) x AAD43514 (1-1344)

QY 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerLys 20
DB 52 ATGTCGGACCTCCACCGCGCGAGCACCAGTCCGCGCCACCGCGCTCCGCGCAAG 111
QY 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrIlyProLeuGlnAlaGlyAspArg 40
DB 112 CTGGCGCCGCTCATCGACGCTCCGCGCTCTTCTACAAGCGCTCCAGCGCGCGACCGT 171
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60
DB 172 GGGGAGCAGCAGGTGCGCTTCTATAGCGGCTTCTCCGCCACCGCGCGCTCCGCGCGC 231
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 232 ATCGGACACCTTCTTCCCGCGTTCCAGCGCAGCGACTCTCCCGCCACCGAGCGCAG 291
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100
DB 292 CCGCGGAGCGCATCTCCACCTCGTCTCGACGACCTCTCCGCGGGGTTTCAGCGGCC 351
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrProProSerSerProGluProTyr 120
DB 352 TGGTTCGACAGATCAAGATCGGCGCCATTCAGTGGCCACGAGTTCGCGGAGCCCTAC 411
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 412 ATCGCAAGTGCTTCGCCAAGGACCGCGGACACCGAGCGTTCGTCTCGGATTCGCGGTC 471
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
DB 472 TCCGGCGTCCGAGTCTCGCGCCCGCGAGCGCGCCGTGTGGCGGACGAGCGCGCGAGGTG 531
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgValSerValAla 180
DB 532 AAGGCCATGGACACCGCGCGGCTCCGCGCGTGTCTCGCGCTACGTTATCCGTTGCC 591
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLeu 200
DB 592 GACGAGGGATGGACTGTGCGCTCGCGCGCGCGGTGTACGGAGGAAAAAGGTGGAGTCTTG 651
QY 201 SerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSer 220
```



Alignment Scores:		Pred. No.:	1,07e-127	Length:	922
Score:			1485.00	Matches:	283
Percent Similarity:			98.27%	Conservative:	1
Best Local Similarity:			97.92%	Mismatches:	5
Query Match:			97.63%	Indels:	0
DB:			6	Gaps:	0
US-10-042-894A-8 (1-289) x AAD43512 (1-922)					
QY	1	MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIlys	20		
DB	53	ATGCCCGACCTCCACCCCGGAGCACCAGTCGCGGCTCCCGCGCTCCCGCAGCAAG	112		
QY	21	LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg	40		
DB	113	CGGGCCCGCTCATCGAGGCTCCGGCTCTTCTCAAGCGCTCCAGCGCGCGACCGT	172		
QY	41	GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg	60		
DB	173	GGGAGCAGAGGTGCTTCTATGAGGGTCTCCGCCCCAGCGCTCCCGCGCGC	232		
QY	61	IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln	80		
DB	233	ATCCGAGACACCTTCTCCCGGTTCCACGCGCAGCTCTCCCGCGAGCGCGAG	292		
QY	81	ProGlyGluProHisProHisValLeuLeuAspLeuAlaGlyPheGlnAlaPro	100		
DB	293	CCGGGAGCGCATCCGACCTCTGCTCGAGACCTCTCCGGGATTTGAGCGGCC	352		
QY	101	CysValAlaAspIleLysIleGlyAlaIleThrProProSerSerProGluProTyr	120		
DB	353	TGCGTCGACAGCATCAAGATCGGCCCATCAGTGGCCACGAGTTCGCGAGCGCTAC	412		
QY	121	IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal	140		
DB	413	ATCGCAAGTGCCTCGCCATGACCGCGGACACAGCGCTCTCTCGGATTCGCGCTC	472		
QY	141	SerGlyValArgValAlaGlyProGluGlyAlaValTyrArgThrGluArgProGluVal	160		
DB	473	TCCGGCTCCGAGTCTGCGCCCGAGGGCGCGTGTGGCGGAGCGAGCGCGAGGTG	532		
QY	161	LysAlaMetAspThrAlaGlyValArgValLeuArgTyrValSerSerValAla	180		
DB	533	AAGGCCATGACACCGCGCGCTCCGCGCGTCTCCGCGCTACGTCATCCGTTGCC	592		
QY	181	AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLeu	200		
DB	593	GACGAGGGGATGAGTGTGCTCGCTCGCGCGGCTACGAGGAGGAAAGGTGGAGTCTTG	652		
QY	201	SerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSer	220		
DB	653	TCACAGCTCGCGAGCTCAAGCGTGTTCGAGGAGCAGACTCTGTTCCACTTCTACTCG	712		
QY	221	AlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaGlyGlyAspGlyGly	240		
DB	713	GGTCTGATTTCTTGGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	772		
QY	241	ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis	260		
DB	773	GTGACGGTGAGCTGTGAGACTTTCCTCATGTGCGCGAGGGGTGATGGGTGATTCACCAC	832		
QY	261	AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr	280		
DB	833	AACCTCTCGGGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	912		
QY	281	ProHisThrGlnProLeuGlyProSer	289		
DB	893	CCTCAGACCGACCTTTGGGTCTCTTCT	919		
RESULT 3					
AAD43512 standard; DNA; 923 BP.					
ID AAD43512					

XX	AAD43512;				
AC	14-NOV-2002	(first entry)			
DT	Maize inositol polyphosphate kinase (IPPK) DNA #2.				
XX	Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;				
XX	nutritional value; animal feed; transgenic; gene; ds.				
KW	Zea mays.				
OS	Key	Location/Qualifiers			
XX	FT	53.736			
FT	CDS	/*tag= a			
FT		/product= "Maize IPPK protein #2"			
XX	WO200259324-A2.				
FN	01-AUG-2002.				
XX	09-JAN-2002; 2002WO-US003120.				
XX	12-JAN-2001; 2001US-0261465P.				
PR	(PION-) PIONEER HI-BRED INT INC.				
XX	Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;				
PI	WPI; 2002-636540/68.				
XX	P-PSDB; AAE26194.				
DR	New inositol polyphosphate kinase polynucleotides and polypeptides,				
PT	useful in modulating phytic acid biosynthesis by decreasing phytate or				
PT	increasing non-phytate phosphorous to improve the nutritional value of				
PT	animal feed.				
XX	Claim 1; Page 60-61; 86pp; English.				
PS	The invention relates to novel inositol polyphosphate kinase (IPPK)				
XX	polypeptides and polynucleotides. Sequences of the invention are useful				
CC	in modulating the phytic acid biosynthesis by decreasing phytate and/or				
CC	increasing non-phytate phosphorous to improve the nutritional value of				
CC	animal feed, or to reduce the environmental impact of animal waste.				
CC	Polynucleotides of the invention are used to produce transgenic plants with an				
CC	altered phenotype. IPPK proteins are used to screen compounds that				
CC	modulate their activity and raising anti-idiotypic antibodies. The				
CC	present sequence is maize IPPK DNA				
XX	Sequence 923 BP; 138 A; 326 C; 286 G; 173 T; 0 U; 0 Other;				
SQ	Alignment Scores:				
	Pred. No.:	4,41e-123	Length:	923	
	Score:	1435.00	Matches:	277	
	Percent Similarity:	96.21%	Conservative:	2	
	Best Local Similarity:	95.52%	Mismatches:	10	
	Query Match:	94.35%	Indels:	1	
	DB:	6	Gaps:	0	
US-10-042-894A-8 (1-289) x AAD43512 (1-923)					
QY	1	MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIlys	20		
DB	53	ATGCCCGACCTCCACCCCGGAGCACCAGTCGCGGCTCCCGCGCTCCCGCAGCAAG	112		
QY	21	LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg	40		
DB	113	CGGGCCCGCTCATCGAGGCTCCGGCTCTTCTCAAGCGCTCCAGCGCGCGACCGT	172		
QY	41	GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg	60		
DB	173	GGGAGCAGAGGTGCTTCTATGAGGGTCTCCGCCCCAGCGCTCCCGCGCGC	232		



Db 562 AAGGCTATGAGACCGTCCGGCGTCCGGCGTCTCCGGCGCTACGTGTCTCCGGCTTCG 621  
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyGlyGlyValle 200  
Db 622 CGACGAGGGATGGACTCGCGCTCCGGCGCGGTGTACGGAGGAAAGGTGGAGTCTT 681  
Qy 200 uSerGlnLeuArgGluLeuLysAlaTyrPheGluGlnThrLeuPheHisPheTyrSe 220  
Db 682 GTACACAGCTGCCGAGCTCAAGCGATGGTGGAGAGCAGACTCTGTCCACTTCTATC 741  
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyGlyGly 240  
Db 742 GGCGTCGATCTTCTGGGCTATGATGCTGCTGCACTCGCAGCGCGGAGGTGGGGTGG 801  
Qy 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValleAspHi 260  
Db 802 GGTAACAGTGAAGCTGGTGACTTTGCCCATGTGGCCGAGGTGATGGGGTGAATGACCA 861  
Qy 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280  
Db 862 CAACCTTCTGGCGAGCTCTGTAGCTGATCAAGTTCGTTTCTGACATTTGTTCCAGAGAC 921  
Qy 280 rProHisThrGlnProLeuGlyProSer 289  
Db 922 TCCTTAGACGCGAGCCTTGGGTCTTCT 949  
RESULT 5  
AAD43522  
ID AAD43522 standard; DNA; 3416 BP.  
AC AAD43522;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #9.  
XX  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; ds.  
XX  
OS Zea mays.  
XX  
FH Key Location/Qualifiers  
FT CDS 72..407  
FT FT /\*tag= a  
FT FT /product= "Maize IPPK protein #6"  
XX  
PN WO200259324-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 09-JAN-2002; 2002WO-US003120.  
XX  
PR 12-JAN-2001; 2001US-0261465P.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX  
DR WPI; 2002-636540/68.  
DR P-PSDB; AAE26201.  
XX  
PT New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX  
PS Claim 1; Page 75-77; 86pp; English.  
XX  
CC The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.

CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA  
XX

SQ Sequence 3416 BP; 893 A; 830 C; 856 G; 837 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1,45e-110 Length: 3416  
Score: 1307.00 Matches: 260  
Percent Similarity: 90.34% Conservative: 2  
Best Local Similarity: 89.66% Mismatches: 9  
Query Match: 85.93% Indels: 20  
DB: 6 Gaps: 1

US-10-042-894A-8 (1-289) x AAD43522 (1-3416)

Qy 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerHisSerLys 20  
Db 72 ATGCCCGACCTCCACCCGCGGAGCACCAAGTCGCGGTCCACCGCGCTCCGCCAGCAAG 131  
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
Db 132 CTGGGCCCACTCATCGACGACTCTGGCCTCTTCTACAAGCCGCTCCAGCCCGGACCGT 191  
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60  
Db 192 GGGGAGCAGAGGTGCGCTTCTATGAGGGGTCTCCGCCCGCGCTCCGCCCGCGC 251  
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
Db 252 ATCCGAGACACCTTCTCCCGGTTCACGCGCAGCGACTCTCCGCCACCGAGGCGCAG 311  
Qy 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100  
Db 312 CCGGGGAGCGGCATCCGACCTCTCTCGACGACCTCTCTCGGGGTTCGAGGCGGCC 371  
Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
Db 372 TCGGTTCGACAGCATCAAGATCGTGCCATCAGTG----- 406  
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 407 -----ACCACGAGCGTTCGTCTCGGATTCCGCGTC 436  
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160  
Db 437 TCCGGCGTCCGAGTCTCGGCCCGCGCGCGGTGTGCGGACGAGCGCGCGGAGGTG 496  
Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180  
Db 497 AAGGCTATGGACATGTGCGCGTCCGCCGCGTCTCCGCCGCTACGTGTCAITCGGTGC 556  
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyGlyGlyValle 200  
Db 557 CGACGAGGGATGGACTCGCGCTCGCGCGCGGTGTACGGAGGAGGAGGAGGAGTCTT 616  
Qy 200 uSerGlnLeuArgGluLeuLysAlaTyrPheGluGlnThrLeuPheHisPheTyrSe 220  
Db 617 GTCACAGCTCGCGAGCTCAAGCGGTGGTTCAGGGGCGAGACTCTCTTTCACCTTCTACTC 676  
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 240  
Db 677 GGCGTCGATCTTCTGGGCTATGATGCTGCTCAGTCCGACGAGCGGAGGTGGGGTGG 736  
Qy 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValleAspHi 260  
Db 737 GGTAAACAGTGAAGCTGTGGACTTTGCCCATGTGGCCGAGGTGATGGGGTGAATGACCA 796  
Qy 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280  
Db 797 CAACCTTCTGGCGGCGCTCTGTAGCTGATCAAGTTCGTTTCTGACATTTGTTCCAGAGAC 856

```

QY 280 rProHisThrGlnProLeuGlyProSer 289
DB 857 TCCTCAGAGCGAGCTTTGGGTCTTCT 884

RESULT 6
ID AAD43518
AC AAD43518;
XX 14-NOV-2002 (first entry)
DE Maize inositol polyphosphate kinase (IPPK) DNA #5.
DE Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 89..424
FT /tag= a
FT /product= "Maize IPPK protein #4"
XX
XX WO200259324-A2.
XX 01-AUG-2002.
XX 09-JAN-2002; 2002WO-US003120.
XX 12-JAN-2001; 2001US-0261465P.
XX (PION-) PIONEER HI-BRED INT INC.
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX P-PSDB; AAE26200.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX Claim 1; Page 73-74; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/ or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is maize IPPK DNA
XX
XX Sequence 899 BP; 140 A; 312 C; 276 G; 171 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.87e-104 Length: 899
Score: 1233.00 Matches: 248
Percent Similarity: 88.42% Conservative: 4
Best Local Similarity: 87.02% Mismatches: 14
Query Match: 81.07% Indels: 20
DB: 6 Gaps: 1

US-10-042-894A-8 (1-289) x AAD43518 (1-899)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
DB 89 ATGCCGACCTCCACCGCGGAGACCAAGTCGCCGGTACCGCGCTCCGCGACCAAG 148
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40

```

```

DB 149 CTGGCCCCACTCATCGAGCGCTCTGGCTCTTCTAAGCGCTCCAGCCCGGACCGT 208
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 209 GGGGAGCAGAGGTGCGCTTCTATAGGCGTTCCTCCGCCACCGCGCTCCCGGCCGC 268
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 269 ATCCGAGACACTTCTTCCCGCGGTTCACCGGACGCGACTCTCTCCACCGAGGCGAG 328
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
DB 329 CCGGGGAGCGCATCTCGTACCTCTCTCGACGACCTCTCTCGGGGTTTGAGGCGCC 388
QY 101 CysValAlaAspIleLeuGlyAlaIleThrTrpProProSerSerProGluProTyr 120
DB 389 TCGTCCGAGACATCAAGATCGTCCATCAGTCACCATG-AGCGAT----- 435
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 436 -----CTGCTCGGATTCCACGTC 453
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
DB 454 TCCGGCGTCCGAGTCTGTCGCCGCCCGCGCGCGTGTGCGGACGCGCGCTGAGGTG 513
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180
DB 514 AAGGCTATGGACATTGTGCGGCTCCCGCGCGTGTCCGGCGTGCATGTCTATCCGTTGC 573
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyVal 200
DB 574 CGCGGAGGGGATGGACTCGCGCTCGCGCGCGGTGTAGCGAGGAGAAAGGTGAGTCTT 633
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSe 220
DB 634 GTCACAGCTGCGCGAGCTCAAGGCGTGTTCGAGGGGCGAGACTCTGTCCACTTCTACTC 693
QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyGlyGly 240
DB 694 GCGCTCGATTCTTCTGGGCTATGATGCTGTGCACTCCGACGAGGCGGAGGTGGGGTGG 753
QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
DB 754 GGTAACAGTGAAGCTGGTGGACCTTGGCCCATGTGGCGGAGGTGATGGGTGATTGACCA 813
QY 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
DB 814 CAACTTCTCGGCGGGCTCTGCTAGTATCAAGTTTGTTCGACATTGTTCAGATTGTCAGAGAC 873
QY 280 rProHisThrGln 284
DB 874 TCCTTAGAGCGCAG 886

RESULT 7
AAD43519
ID AAD43519 standard; DNA; 643 BP.
XX
XX AAD43519;
XX
XX 14-NOV-2002 (first entry)
XX
XX Maize inositol polyphosphate kinase (IPPK) DNA #6.
XX
XX Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
XX nutritional value; animal feed; transgenic; ds.
XX
XX Zea mays.
XX
XX WO200259324-A2.
XX
XX 01-AUG-2002.

```

```
XX 09-JAN-2002; 2002WO-US003120.
PF 12-JAN-2001; 2001US-0261465P.
PR (PION-) PIONEER HI-BRED INT INC.
PA Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
PI WPI; 2002-636540/68.
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
PS Claim 1; Page 74; 86pp; English.
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/ or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is maize IPPK DNA
XX
SQ Sequence 643 BP; 92 A; 261 C; 178 G; 103 T; 0 U; 9 Other;

Alignment Scores:
Pred. No.: 1.85e-65 Length: 643
Score: 808.50 Matches: 162
Percent Similarity: 93.14% Conservative: 1
Best Local Similarity: 92.57% Mismatches: 10
Query Match: 53.16% Indels: 3
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AAD43519 (1-643)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIys 20
Db 117 ATGCCCGACCTCCACCCCGGAGACCAAGTCGCGCTACCCGCGCTCCGCCAGCAAG 176
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 177 CTGGCCCGCTCATCGAGGCTCCGGCTCTTCTACAGCGCTCCAGCGCGGACCGT 236
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 237 GGGAGACGAGGTCGCTTCTATGAGCGTCTCTCCGCCACCGCGCGTCCCGCGCGC 296
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 297 ATCCGAGACACCTCTTCCCGGTTCCAGCGACCGGACTCTCCCCAGCGCGGAG 356
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 357 CCGCGGAGCGGCATCCGACCTCTCTCCGACGACTCTCTCCGGGTTTGAGSGCGCC 416
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120
Db 417 TCGGTGCGAGACATCAAGATCGCGCCATCATCGTGGCCACCGAGTTCGCGGAGCCCTAC 476
QY 121 IleAlaIleCysLeuAlaIleAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 477 ATCGNCAAGTACTTNGCCAAAGACCGCGGACCAAGAGGTTCTCTCGGATTCGCGGT 536
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArg-ProGluVal 160
Db 537 TTGC--GTCCGAGTGTGCGGCCCGGAGGCGCGGTGTGGCGAGCGGAGCGCGCGGGGT 594
QY 160 IlyAlaMetAspThr-AlaGlyValArgValLeuArg 173
```

```
Db 595 GAANGCTATGACACCCGTCGGNGNCCGGCGNGTGTCTCGG 635
RESULT 8
RAD43517
ID AAD43517 standard; DNA; 1020 BP.
XX AAD43517;
XX 14-NOV-2002 (first entry)
XX P. argentatum inositol polyphosphate kinase (IPPK) DNA #1.
XX Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
XX nutritional value; animal feed; gene; ds.
XX Parthenium argentatum.
XX
FH Key Location/Qualifiers
FT CDS 21..908
FT /*tag= a
FT /product= "P. argentatum IPPK protein #1"
XX
PN WO200259324-A2.
XX 01-AUG-2002.
XX 09-JAN-2002; 2002WO-US003120.
XX 12-JAN-2001; 2001US-0261465P.
XX (PION-) PIONEER HI-BRED INT INC.
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX P-PSDB; AAE26199.
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX Claim 1; Page 71-72; 86pp; English.
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/ or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is P. argentatum IPPK DNA
XX
SQ Sequence 1020 BP; 297 A; 195 C; 219 G; 309 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.04e-59 Length: 1020
Score: 742.50 Matches: 149
Percent Similarity: 67.93% Conservative: 45
Best Local Similarity: 52.10% Mismatches: 79
Query Match: 48.82% Indels: 13
DB: 6 Gaps: 5

US-10-042-894A-8 (1-289) x AAD43517 (1-1020)
QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIysLeuGlyPro 23
Db 24 CTCAGGGCCCCAGATCATCATCGTGTGACATGAAGCTGGGCTCGGAGAGCTTGGCCCA 83
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
```

```
Db 84 CTCATTGATGATTAGCGCGGTTTTTACAAACCACTGCAGGGTGATVAAACCGTGGGTGAGAA 143
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
Db 144 GAAGTAGCCTTTTATGAATCATTTCTCTCAATATATATTCAGAACACATATGCG--- 200
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 201 AAATTCCTTCTATATATTATGCGCAAAATCATG-----AAGGCATCCACTGGCTCT 254
QY 84 ProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103
Db 255 GACCATCTCTCATCATGCTGTGCAAGATCTTATCATCAGCTCATGTCAACCCATCTGTAAAG 314
QY 104 AspIleLeuIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAlaLys 123
Db 315 GACATCAAAATCGGTCAGACATATGGCGCCAGAGCTTCGAGCGGTACATTCGAAA 374
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143
Db 375 TGCTTAAAAAGGATAGGGAAGACACAGTATTCATTCGGATTCAGAGTCTCCGGGTG 434
QY 144 ArgValValGlyProGluGlyAla---ValTrpArgThrGluArgProGluValLysAla 162
Db 435 CAAGTCTATATCATGATATGGTCAGCGGTTTTATAGCCCTCATAGAAATACATGCGTAAA 494
QY 163 MetAspThrAlaGlyValArgValLeuArgTyrValSerSerValAlaAspGlu 182
Db 495 ACCGGCCAGCTCATCTAGACTACTCTTAGGAATTTGTTCTTCTTAACCCGCTCGCA 554
QY 183 -----GlyMetAspCysAlaLeuAlaAlaValTyrGly 194
Db 555 GAGATGGAATGCGCACAGCGCTAGCGCGGATGTTCTTTAGCATCTTTTGTATGTT 614
QY 195 GlyGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThr 214
Db 615 GGCCCTAATGGATATAGTCACTCACTGATGGAATGGAACATGTTTGAAGATCAACA 674
QY 215 LeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAla 234
Db 675 ATTTACCACCTCTATGCTTTCTTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 734
QY 235 GlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGly 254
Db 735 GGTGCT---CGGTCAACACGACAGAGTCAAACTTATTTGATTTTCTCATGTACAGATGT 791
QY 255 AspGlyValIleAspHisPheLeuGlyGlyLeuCysSerIleLeuLysPheValSer 274
Db 792 AATGGTGTTATGATCACAAATTTCTTGGGTGGGCTCTGTTCTTTGATTAAGTTCTTCT 851
QY 275 AspIleValProGluThr 280
Db 852 GACATACTTTCGGAGACA 869
RESULT 9
ID AAD43516
XX
AC AAD43516;
XX
XX 14-NOV-2002 (first entry)
XX Eucalyptus grandis inositol polyphosphate kinase (IPPK) DNA.
DE DE
XX Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
XX nutritional value; animal feed; gene; ds.
XX Eucalyptus grandis.
OS
XX Key Location/Qualifiers
FH 116..1048
FT CDS /*tag= a
FT /product= "E. grandis IPPK protein"
```

```
XX WO200259324-A2.
FN
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US0003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
PI WPI; 2002-636540/68.
XX P-PSDB; AAE26198.
DR
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
XX Claim 1; Page 68-70; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are used to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is Eucalyptus grandis IPPK DNA
XX
XX Sequence 1195 BP; 342 A; 240 C; 294 G; 319 T; 0 U; 0 Other;
```

## Alignment Scores:

Pred. No.:	3 456-57	Length:	1195
Score:	722.50	Matches:	149
Percent Similarity:	66.31%	Conservative:	38
Best Local Similarity:	52.84%	Mismatches:	84
Query Match:	47.11	Indels:	6
DB:	6	Gaps:	6

US-10-042-894A-8 (1-289) x AAD43516 (1-1195)

```
QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
Db 119 CTCAAGGTCCTCGGATCATCAAGTCGCGGTCAACCGGGAGACGGGGGAAAGCTGGGCCA 178
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 179 CTGGTGGATGATTCGGGCGCTTCTATAGCTCTCCAGAGCGATCATCGGAGACAG 238
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
Db 239 GAAGTGGCTTTTACGAGTCATTCTATCCATCCGAGATCCAGGTCCACATTCGC--- 295
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 296 AAATTCCTTCTCTCGGTTTCAGGAACTAAGACTATT-----GAGCGCTCTGTATGATCG 349
QY 84 ---ProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysVal 102
Db 350 GGTCTCTCAACCTCACCTGTTCTTGAGGATCTCGTCTCGGTCGACGACCATCTCTC 409
QY 103 AlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAla 122
Db 410 ATCGACATCAAGCTGATCCAGAACATGTTATCGAGAGCTCTCCAGAGTACATCCAA 469
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
Db 470 AAGTGCTTAGAGAAAGATCGAAATAGCAACAGCGTTTCTATGGGTTTATGAGTTTCTGG 529
```





PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147203P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	14-OCT-1999;	99US-0159638P.			
PR	18-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160770P.			

Alignment Scores:

Pred. No.: 1,76e-56 Length: 1130  
 Score: 714.50 Matches: 150  
 Percent Similarity: 66.55% Conservative: 39  
 Best Local Similarity: 52.82% Mismatches: 82  
 Query Match: 46.98% Indels: 13  
 DB: 3 Gaps: 8

US-10-042-894A-8 (1-289) x AAC48750 (1-1130)

QY	4	LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIysLeuGlyPro	23
DB	94	CTCAAGGTCCTCGAACACCAAGTTGCTGTCACATTGCTAGTGGGAGGCTCGGTCCA	153
QY	24	LeuIleAspGlySerGlyLeuPheTyrIysProLeuGlnAlaGlyAspArgGlyGluHis	43
DB	154	CTCGTAGATGACCAAGCGGTTCTTCAAGCCACTTCAGGGAGATTCTCTGGCGAACAC	213
QY	44	GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArgAsp	63
DB	214	GAGGCTAAGTTCTATGAGTCTTTCATCGAACATGAGGTTCCAGATCACATCCAT---	270
QY	64	ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu	83
DB	271	AGATACTTCCCGGTGTATCAGCGCACTCAGCTAGTT-----GAAGCATCTGATGATCT	324
QY	84	ProHis--ProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysVal	102
DB	325	GGCAAGCTTCTCATCTGTTCTTGTATGATGTTGTTTCAGGTCAGCAACCCGTCGGTA	384
QY	103	AlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAla	122
DB	385	ATGGATGTTAAGATTGGATCTAGACATGTCACCGGATGTATCAGAAGAACTTCAAG	444
QY	123	LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly	142
DB	445	AAATGATTAAAGAAAGATAGACAGACCACCGTTTCGTTGGGTTTCAAGGTTTCAGGT	504
QY	143	ValArgValValGly---ProGluGlyAlaValTrpArgThrGluArgProGluValLys	161
DB	505	TTTAAGATTTTGTATCACCACCAAGATCAAGTTTGGAGAGCTGAGAAGAGCTTGTCTT	564
QY	162	AlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSer-----SerVal	179
DB	565	GGGTATAATGTCAGATGGTCTAGATTGGCTCTGAGAGAGTTTGTGTCATCGAACTCTCC	624
QY	180	AlaAspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGly	197
DB	625	GCTGACTCTAACTTGACCAACACTGTCTTTTGTGATCAGAGTTTATGGCGTTTAAAC	684
QY	198	GlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHis	217
DB	685	GGGATCTTAGCGCAGTTTGTGGAGCTTAAAGATTGTTTCGAAACCAACCGCTTACCAT	744
QY	218	PheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaVal---AlaAlaGlyGly	236
DB	745	TTCAATTCTCGATTCTGTATGATTTTACGAGAAATGAATCAATCTTGTGCAAGGAGA	804

Qy 237 AspGlyGly-----GlyValThrVallysLeuValAspPheAlaHisValAlaGlu 253  
|||  
Db 805 GATGATGCACCGGCACACCGGCACAAAGTGAAGCTGGTGGATTTCGCCGCAATGTTCTTGAT 864  
|||  
Qy 254 GlyAspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheVal 273  
|||  
Db 865 GGAAACGGGTGTCATCGACCAATAATTCTTGGTGGACTCTGCTCTTTTCATAAAGTTTCATC 924  
|||  
Qy 274 SerAspIleVal 277  
|||  
Db 925 AAAGATATTTCTT 936  
|||  
RESULT 11  
AAC40313  
ID AAC40313 standard; DNA; 1130 BP.  
XX  
AC AAC40313;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 27858.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.





```

PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153707P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 8,41e-53
Score: 674.50
Percent Similarity: 65.36%
Best Local Similarity: 50.36%
Query Match: 44.35%
DB: 3

US-10-042-894A-8 (1-289) x AAC33685 (1-1104)

QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerLysLeuGlyPro 23
Db 132 CTCAAGTCCTGACATCAGTTGTCAGGACATTCCTTAACGCGGAGCTGTCT 191

QY 24 LeuLeuAspGlySerGlyLeuPheTyrLeuProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 192 CTCGTAGATCAAGGTCGTCTTCAAGCCACTTCAGGCGCATTCCTCGTGTGAATC 251

QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
Db 252 GAGGTAAGTTCTACGAATCTTTCTCTCAACACACAGAGGTTCCAGAACACATCCAT--- 308

```

```

QY 64 ThrPhePheArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 309 AGATATTTCCCGTGTATCAGCGCACTCAAGCAGTT-----GAAGTTCTGTATGAGCA 362

QY 84 ProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103
Db 363 -----GCCATCATGTGTGGAAAAATCTTTGCAGAAATACTCAAAACCATCAGTAATG 416

QY 104 AspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAlaLys 123
Db 417 GATGTTAAGATGGTTTCGAGAAACATGTCATCTCTGATCTGAGAAATACATCCAAAAA 476

QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143
Db 477 TGTTTGAAGAAGACACGGGTACCAACACCGTGTATCGGGTTTCAGGATCTCTGGTTTC 536

QY 144 ArgValValGly---ProGluGlyAlaValTrpArgThrGluArgProGluValLysAla 162
Db 537 GAAATGTATGATCACAAGAAATCGAGTTTCTGGAAGCCCGAGAGGAAGCTTCTTCGCGG 596

QY 163 MetAspThrAlaGlyValArgValLeuArgValTyrValSer-----SerValAla 180
Db 597 CTCGATGTAGTGGAGCGGAGATTGACTCTGAGGAAGTTTGTATCATCTAACTCACTTCG 656

QY 181 AspGluGlyMet-----AspCysAlaLeuAlaAlaAlaValTyrGlyLysGlyGly 198
Db 657 GACACTGGCTCGAAACCTGACTCTGCTTTGCTCGAGTGTTCACGCGGTTCGCCACGGG 716

QY 199 ValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPhe 218
Db 717 ATCTTAACGCGATGCTGGAACCTCAAGACCTGGTTCGAGAACCAACGCTCTACCAATTC 776

QY 219 TyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaGlyGlyAspGly 238
Db 777 AACTCTTGTTCGATTTTAATGCTCTATGAGAAATGATCCATCTTGNAGGGAATGATGAT 836

QY 239 GlyGly---ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
Db 837 GATGCTAGACCACCAAGTCAAGCTGGTGGATTTTGCTCATGTTCTTGATGTAATGCTGTC 896

QY 258 IleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
Db 897 ATTGACCAATACTCTTGGTGTCTTCTGCTCTTCTTCAATAAATCTTATCGTGAGATTCTT 956

RESULT 13
AAC39023
ID AAC39023 standard; DNA; 1168 BP.
XX AAC39023;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 23077.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.

```

PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0134449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147415P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149923P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 27-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153700P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-015753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.





PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 30-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149829P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158023P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.

PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

## Alignment Scores:

Pred. No.: 9,74e-53 Length: 1243  
Score: 674.50 Matches: 141  
Percent Similarity: 65.36% Conservative: 42  
Best Local Similarity: 50.34% Mismatches: 86  
Query Match: 44.35% Indels: 11  
DB: 3 Gaps: 7

US-10-042-894A-8 (1-289) x AAC38692 (1-1243)

QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23  
DB 269 CTCAAAGTCCCTGAACATCAGGTTCGAGGACATCTGCTAAAGCGGAGCCTGCTCT 328  
QY 24 LeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43  
DB 329 CTCGTAGATGACAGAGGTTCGTTCTCAAGCCACTTCAGGGGATTCCTGCTGTAATC 388  
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgLleArgasp 63  
DB 389 GAGGTAAGTTCTACGAATCTTCTCTCTCAACACAGAGGTTCCAGAACACATCCAT 445  
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83  
DB 446 AGATATTTCCCGGTGTATCCGGCACTCAAGCAGTT-----GAAGTTCTCATGGACA 499  
QY 84 ProHisProHisValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103  
DB 500 -----GCCATCATGCTGTGGAATACTCTTCGAGATACCTCAAAACCATCATGTAATG 553  
QY 104 AspLleLysLleGlyAlaLleThrProProSerSerProGluProTyrLleAlaLys 123  
DB 554 GATGTTAAGTGGTTCGAGAACTGGTATCCTGATCATCTGGAAGATACATCAAA 613  
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143  
DB 614 TGTGTTGAGAAAGACACGGGTACCAACACCGGTGATCGGGTTTCAAGATCTCTGTTTC 673  
QY 144 ArgValValGly-----ProGlyAlaValThrPheArgThrGluArgProGluValLysAla 162  
DB 674 GAAAGTGTATGATCACAAAGAAATCGAGTTCTGGAAGCCGAGAGGAAGCTTCTCGCGG 733  
QY 163 MetAspThrAlaGlyValArgArgValLeuArgArgTyrValSer-----SerValAla 180  
DB 734 CTCGATGTAGTGGAGCGGAGATGTGACTCTGAGGAAGTTGTATCATCTAACTCATCTTCG 793  
QY 181 AspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyLysGlyGly 198  
DB 794 GACACTGGCTCGAAACCTGACTGCTGTTGCTCGAGTGTTCACGGCGTTCCACCGG 853  
QY 199 ValLeuSerGlnLeuArgGlnLeuLysAlaThrPheGluGlnThrLeuPheHisPhe 218  
DB 854 ATCTTAACGACGTTGCTGGAACCTCAAGACCTGTTGAGAACCAACCGCTCTACCATTC 913  
QY 219 TyrSerAlaSerLleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyLysGly 238  
DB 914 AACTCTGTGCTGATTTTATGGTCTATGAGATGAATCCATCTTGAAGGAATGATGAT 973  
QY 239 GlyGly-----ValThrValLysLeuValAspPheAlaHisValAlaGlyAspGlyVal 257  
DB 974 GATGCTAGACCACAAAGTCAAGCTGGTGGATTTTGCTCATGTTCTTGTGTAATGCTGTC 1033  
QY 258 IleAspHisAsnPheLeuGlyGlyLeuCysSerLeuLleLysPheValSerAspLleVal 277  
DB 1034 ATTGACCAATACTTCTTGGGTGGTCTTGTCTCTTTCATATAAATTCATCTGATGATCTT 1093

RESULT 15  
AAD43515

ID AAD43515 standard; DNA; 1105 BP.

XX AC AAD43515;

XX 14-NOV-2002 (first entry)

XX Soybean inositol polyphosphate kinase (IPPK) DNA.

XX Soybean; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
XX nutritional value; animal feed; transgenic; gene; ds.

XX Glycine max.

XX Key Location/Qualifiers  
XX CDS 12..851  
XX /\*tag= a  
XX /product= "Soybean IPPK protein"

XX WO200259324-A2.

XX 01-AUG-2002.

XX 09-JAN-2002; 2002WO-US003120.

XX 12-JAN-2001; 2001US-0261465P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;

XX WPI; 2002-636540/68.

XX P-PSDB; AAE26197.

XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
XX useful in modulating phytic acid biosynthesis by decreasing phytate or  
XX increasing non-phytate phosphorous to improve the nutritional value of  
XX animal feed.

XX Claim 1; Page 66-67; 86pp; English.

XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
XX polypeptides and polynucleotides. Sequences of the invention are useful  
XX in modulating the phytic acid biosynthesis by decreasing phytate and/or  
XX increasing non-phytate phosphorous to improve the nutritional value of  
XX animal feed, or to reduce the environmental impact of animal waste.  
XX Polynucleotides of the invention are used to produce transgenic plants with an  
XX altered phenotype. IPPK proteins are used to screen compounds that  
XX modulate their activity and raising anti-idiotypic antibodies. The  
XX present sequence is soybean IPPK DNA

XX Sequence 1105 BP; 279 A; 311 C; 233 G; 282 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 4.94e-50 Length: 1105  
Score: 644.50 Matches: 141  
Percent Similarity: 62.72% Conservative: 34  
Best Local Similarity: 50.54% Mismatches: 85  
Query Match: 42.37% Indels: 19  
DB: 6 Gaps: 8

US-10-042-894A-8 (1-289) x AAD43515 (1-1105)

QY 7 ProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyProLeuLleasp 26

DB 24 CCGAGACACCAAGTGGCGGGGCAACAGGCAAGGCAAGCAATCTCTGGGCCCTCTGTCGAC 83

QY 27 GlySerGlyLeuPheTyrLysProLeuGlnAlaGly-----AspArgGlyGlu 42

DB 84 GATTTTGAAAATTTCTACAGCCCTCCAGACCAACAAAGACGACGACCCCGGCTCC 143

QY 43 HisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgLleArg 62

DB 144 ACCGAACTCTCTTTTACACCTCTCTCGCGCC-----GCGGCCACGACTACTCTCCATCGC 200

```
QY      63  AspThrPheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGly  82
Db      ::::|||||
201  ---TCTTCTTCCCGCCCTTTACGGCACCCGCTCTCTGGACGCTCTCCGACGGCTCCGGT  257
QY      83  GluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaProCysVal  102
Db      ::::|||||
258  ---CCCCACCTCACCCTGGTCTCTGGAGGACCTCTCTGGGCTACTCCAAACCCCTCCGTC  314
QY     103  AlaAspIleLeuGlyAlaIleThrTrpProSerSerProGluProTyrIleAla  122
Db      ::::|||||
315  ATGGACGTAAGATCGGCTCCAGACCTGGACCTGGGAGACTCCGAGGACTACATCTGC  374
QY     123  LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly  142
Db      ::::|||||
375  AAGTGCTCAAGAGGACAGAGAGCTCTCTAGCTTGCCTTGGGTTTCAGAAATCTCGGA  434
QY     143  ValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluValLysAla  162
Db      ::::|||||
435  GTCAAG-----GACTCTATCTCTCTCTGGGACCTACGAGAAATCTCTCCAGTGT  485
QY     163  MetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-----  178
Db      ::::|||||
486  CTATCCGCCCATGCTGTCTCTCAACAAGTTCTCTTAATAATATCAAC  545
QY     179  ValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGly  198
Db      ::::|||||
546  CATGATGATCATCATCCGATTGCGCTTTCGCAACGAGGTCTAC-----GCGGCC  596
QY     199  ValLeuSerGlnLeuArgGluLeuLysAlaTyrPheGluGluGlnThrLeuPheHisPhe  218
Db      ::::|||||
597  GTTTGGAGCGCTTCAGAGCTCAGGACTGGTTGAGGTTTCAGACGGTGTATCCTTC  656
QY     219  TyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGly  238
Db      ::::|||||
657  TATCTCTGTCTCTTCTTGTGTGTACGAGAGGAT-----CTAGGGAAGGGAAGCT  710
QY     239  GlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIle  258
Db      ::::|||||
711  ACCAACCTCTGGTCAAACTCGTTGACTTTGCACACGTTGGTGAGCGAAACGGTGTCT  770
QY     259  AspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal  277
Db      ::::|||||
771  GATCACAACTTCTGGTGGCTTTGTCTCTCATCAAGTTCCTCAAGGATATCCTA  827
```

Search completed: March 27, 2004, 06:19:15  
Job time : 494 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 27, 2004, 04:39:30 ; Search time 4448 Seconds  
(without alignments)  
2816.128 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVAGHRASASK.....IKFVSDIVPETPHTQPLGPS 289

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US10042894/runat\_24032004.145047.16946/app.query.fasta\_1.455  
-DB=GenEmbl -OPMT=fastap -SUFFIX=xge -MINMATCH=0.1 -LOOPCT=0 -LOAPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10042894.cgn 1 1 3731 @runat\_24032004.145047.16946 -NCPU=6 -ICPU=3  
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vl:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

29: em.vi:\*  
30: em.htg.hum:\*  
31: em.htg.inv:\*  
32: em.htg.other:\*  
33: em.htg.mus:\*  
34: em.htg.pln:\*  
35: em.htg.rod:\*  
36: em.htg.mam:\*  
37: em.htg.vrt:\*  
38: em.sy:\*  
39: em.htg.hum:\*  
40: em.htg.mus:\*  
41: em.htg.other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1521	100.0	1344	6	AX513570 Sequence
2	1485	97.6	923	6	AX513568 Sequence
3	1435	94.3	923	6	AX513566 Sequence
4	1406.5	92.5	1169	6	AX513564 Sequence
5	1307	85.9	3416	6	AX513583 Sequence
6	1233	81.1	899	6	AX513578 Sequence
7	1148.5	75.5	132170	8	AP005749 Oryza sat
8	1144.5	75.2	1570	8	AK072296 Oryza sat
9	808.5	53.2	643	6	AX513580 Sequence
10	742.5	48.8	1020	6	AX513576 Sequence
11	722.5	47.5	1195	6	AX513574 Sequence
12	718.5	47.2	25054	8	LUS310150 Linum usi
13	714.5	47.0	903	8	ATH243592 Arabidops
14	714.5	47.0	903	8	AY072621 Arabidops
15	714.5	47.0	903	8	AY147936 Arabidops
16	714.5	47.0	1166	8	AF412073 Arabidops
17	714.5	47.0	1170	8	ATH245521 Arabidops
18	714.5	47.0	57246	8	AB010069 Arabidops
19	704.5	46.3	1130	8	AY085862 Arabidops
20	674.5	44.3	1243	8	AY087217 Arabidops
21	673.5	44.3	1390	8	AY136378 Arabidops
22	673.5	44.3	116763	8	ATTT211 Arabidops
23	664	43.7	115040	8	AC124954 Medicago
24	658.5	43.3	861	8	ATH404678 Arabidops
25	658.5	43.3	861	8	AY147935 Arabidops
26	648	42.6	129248	2	AC127018 Medicago
27	644.5	42.4	1105	6	AX513572 Sequence
28	469.5	30.9	43428	8	AP004981 Lotus cor
29	464	30.5	519	6	AX513581 Sequence
30	340.5	22.4	116893	8	AP004772 Oryza sat
31	329	21.6	353	6	AX513582 Sequence
32	288	18.9	85095	3	AC004573 Drosophil
33	288	18.9	142257	2	AC017491 Drosophil
34	288	18.9	182726	3	AC008002 Drosophil
35	288	18.9	307741	3	AE003589 Drosophil
36	286	18.8	1219	3	AY119207 Drosophil
37	217.5	14.3	3003	3	AK116726 Ciona int
38	201.5	13.2	1251	9	AF432853 Homo sapi
39	199.5	13.1	5432	10	BC052463 Mus muscu
40	194	12.8	5432	10	AY014898 Rattus no
41	193	12.7	1341	6	AR380264 Sequence
42	192	12.6	2391	6	AX816462 Sequence
43	192	12.6	2391	9	AF393812 Homo sapi
44	192	12.6	2890	6	AX504248 Sequence
45	192	12.6	2890	6	AX504248 Sequence

# ALIGNMENTS

RESULT 1

```

AX513570
LOCUS AX513570 1344 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 7 from Patent WO02059324.
ACCESSION AX513570
VERSION AX513570.1 GI:23559670
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 7 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source
Location/Qualifiers
1..1344
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
52..921
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52954.1"
/db_xref="GI:23559671"
/db_xref="REMBL:CAD52954"
/translation="MSDLHPPEHQVAGHRASAKLPLIDSGLFYKPLQADGRGEHE
VAFYFAFAHAAPARIDTFPPRHGRLLPTEAQGPCEPHPLVLDLLAGFAPCV
ADIKGATWPPSSPEPIAKLMDRGTTSLVILGFRVSVGVVPEGAVWRTERPEV
KAMDAGVRVLRVSVVADBGMDALAAAVYGGKGVLSQLRELKAWFEQTLFHF
YSASILLGYDAAVAAGDGGGVTVKLVDFAHVAGDGVIDHNFJLGLCSLKFVSDI
VPETPTQPLGPS"
ORIGIN
Alignment Scores:
Pred. No.: 6.98e-106 Length: 1344
Score: 1521.00 Matches: 289
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-042-894A-8 (1-289) x AX513570 (1-1344)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
DB 52 ATGTCGACCTCCACCGCGGAGCACCACCAAGTCGCGGCCACCGCGCTCCGCCGCAAG 111
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 112 CTGGGCGCGCTCATCGCGCTCGCGCTCTCTCAAGCCGCTCCAGCGCGCGACCGT 171
QY 41 GlyGluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 172 GGGAGCAGAGTGCCTTCTATGAGCGTCTTCGCGCCACCGCGCTCCGCCCGC 231
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 232 ATCCGAGACACCTCTTCGCCCGGTTCCACGCGCACGCGACTCTCCCCACCGAGCGCAG 291
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
DB 292 CCGGGAGCGCGATCCTCACCCTGCTCGACGACTCTCTCGGGGTTTCAGGCGCC 351
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTriProProSerSerProGluProTyr 120
DB 352 TGGCTCGCAGACATCAAGTCGGCGCCATCAGTCGCCACCGAGTTTCGCGGAGCCCTAC 411
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 412 ATGCCAAGTGCCTCGCCAGGACCGCGGACCACGAGCGTTCTGCTCGGATTCGCGCTC 471

```

```

QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
DB 472 TCCGCGCTCCAGTCTCGCGCCCGAGCGCCCGCTGTGGCGAGCGGAGCGCGGAGTG 531
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgfyrValSerSerValAla 180
DB 532 AAGGCCATGGACACCGCGCGCTCCGCGCGCTGCTCCGCGCTACGTCTCCGTTGCC 591
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyGlyGlyValLeu 200
DB 592 GACGAGGGGATGGACTGTGCGCTCGCGCGCGGTGTACGAGGAAAGGTGGAGTCTTG 651
QY 201 SerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSer 220
DB 652 TCACAGCTCGCGAGCTCAAGCGGTTCGAGGAGCAGACTCTGTCCACTTCTACTCG 711
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 240
DB 712 CGGTCGATTCCTCTGGGCTATGATGCTGCTCAGTCGCGAGCGGAGATGGGGTGGG 771
QY 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260
DB 772 GTGACGGTGAAGCTGGTGGACTTTGCCCATGTGCGCGAGGTGATGGGTGATTCACCC 831
QY 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280
DB 832 AACTTCCTGGCGCGGCTCTGCTCGTATCAAGTTCGTTCTGACATTTGTTCCGAGACT 891
QY 281 ProHisThrGlnProLeuGlyProSer 289
DB 892 CCTCATACGACGCTTTGGGTCTTCT 918
RESULT 2
AX513568
LOCUS AX513568 923 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 5 from Patent WO02059324.
ACCESSION AX513568
VERSION AX513568.1 GI:23559668
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 5 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source
Location/Qualifiers
1..923
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
53..922
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52953.1"
/db_xref="GI:23559669"
/db_xref="REMBL:CAD52953"
/translation="MPDLHPPEHQVAGHRASAKPLIDSGLFYKPLQADGRGEHE
VAFYFAFAHAAPARIDTFPPRHGRLLPTEAQGPCEPHPLVLDLLAGFAPCV
ADIKGATWPPSSPEPIAKLMDRGTTSLVILGFRVSVGVVPEGAVWRTERPEV
KAMDAGVRVLRVSVVADBGMDALAAAVYGGKGVLSQLRELKAWFEQTLFHF
YSASILLGYDAAVAAGDGGGVTVKLVDFAHVAGDGVIDHNFJLGLCSLKFVSDI
VPETPTQPLGPS"
ORIGIN
Alignment Scores:
Pred. No.: 2.35e-103 Length: 923
Score: 1485.00 Matches: 283
Percent Similarity: 98.27% Conservative: 1
Best Local Similarity: 97.92% Mismatches: 5

```

```
Query Match: 97.63% Indels: 0
DB: 6 Gaps: 0
US-10-042-894A-8 (1-289) x AX513568 (1-923)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerHisSerHis 20
DB 53 ATGCCGCGACCTCCACCGCGGAGACCAAGTCGCGGTACCGCGCTCCGCCAGCAG 112
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 113 CCGGCGCGCTCATCGACGCTCCGGCTCTCTACAGCCGCTCCAGCCGCGCGCT 172
QY 41 GlyGluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60
DB 173 GCGGAGACGAGTGCCTTCTATGAGCGTCTCCGCCAGCGCTCCGCCGCGCGCT 232
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 233 ATCCGAGACACCTTCTCCCGCGGTTCACGCGACCGACTCTCCACCGCGCGCG 292
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
DB 293 CCGCGGAGCGCGACCTCCGCACTCTCTCGACGACCTCTCCGCGGATTTGAGCGCGC 352
QY 101 CysValAlaAspIleGlyIleGlyAlaIleThrProProSerSerProGluProTyr 120
DB 353 TCGGTCCGAGACATCAAGATCGCGCTCATCGTCCGCGACCGAGTTCCGCGAGCCT 412
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 413 ATCCGCAAGTGCCTCGCATGACCGCGGACCGCGGCTCTCTCGGATTTCCGCGTC 472
QY 141 SerGlyValArgValGlyProGluGlyAlaValAlaValTyrArgThrGluArgProGluVal 160
DB 473 TCCGCGTCCGAGTCTCGGCCCGCGCGCGCTGTGGCGGACGCGCGCGAGGTG 532
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAla 180
DB 533 AAGGCATGAGACACCGCGCGCTCCGCGCGTGTCTCCGCGTACGTGTCTATCGTTC 592
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyGlyGlyValLeu 200
DB 593 GACGAGGGGATGGATGTGCGCTCGCGCGCGGTGTACGAGGAGAAAGGTGGAGTCT 652
QY 201 SerGlnLeuArgGlyLeuLysAlaTyrPheGluGluGlnThrLeuPheHisPheTyrSer 220
DB 653 TCACAGCTCGCGGAGCTCAAGCGGTGTTCGAGGAGCAGACTCTGTCTTCTACTCG 712
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240
DB 713 GCGTCAATCTTCTGGGCTATGATCTGCTGCACTCGAGTCGAGGAGCGAGGTGGG 772
QY 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260
DB 773 GTGACGGTGAAGCTGTGGACTTTCCCATGTCGCGGAGGAGTGAAGGAGTGAAGT 832
QY 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280
DB 833 AACTCTCTGGGCGGCTCTGCTCGCTGATCAAGTCTGCTTTCTGACATTTGTTCCAG 892
QY 281 ProHisThrGlnProLeuGlyProSer 289
DB 893 CCTCAGACGACGCTTTGGGTCTCTTCT 919

RESULT 3
AX513566
LOCUS AX513566 923 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 3 from Patent WO02059324.
ACCESSION AX513566
VERSION AX513566.1 GI:23559666
KEYWORDS Zea mays
```

```
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Caboon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 3 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
Location/Qualifiers
source 1..923
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
53..736
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52952.1"
/db_xref="GI:23559667"
/translation="MPDLHPPEHGVAGHRASAKPGPLIDGSLFKYKPLQAGDRGEHE
VAFYEAFSAHAAVPARIDTFFRPHGTRLLPTBAQCEPHPLVLDLLAGFEAPCV
ADIKGALTWPSPSPPEPIAKLAMDGTTSLVIGRVSQVYVVEGAVWRTSEPEV
KAMDTGVRRVLRIRYVSACRRGDLRGGRVRRKWSLVTAARAQGVVRGAASVPL
LLGVDSGL"

ORIGIN
Alignment Scores:
Pred. No.: 1,4e-99 Length: 923
Score: 1435.00 Matches: 277
Percent Similarity: 96.21% Conservative: 2
Best Local Similarity: 95.52% Mismatches: 10
Query Match: 94.35% Indels: 1
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AX513566 (1-923)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerHisSerHis 20
DB 53 ATGCCGCGACCTCCACCGCGGAGACCAAGTCGCGGTACCGCGCTCCGCCAGCAG 112
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 113 CCGGCGCGCTCATCGACGCTCCGGCTCTCTACAGCCGCTCCAGCCGCGCGCT 172
QY 41 GlyGluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60
DB 173 GCGGAGACGAGGTGCGCTTCTATGAGCGGTTCGCGCGCGCGCTCCGCCGCGCG 232
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 233 ATCCGAGACACCTTCTCCCGCGGTTCACGCGACCGACTCTCTCCACCGAGCGCG 292
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
DB 293 CCGCGGAGCGCGACCTCCGCACTCTCTCGACGACCTCTCCGCGGATTTGAGCGCGC 352
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrProProSerSerProGluProTyr 120
DB 353 TCGGTCCGAGACATCAAGATCGCGCTCATCGTCCGCGACCGAGTTCCGCGAGCCT 412
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 413 ATCCGCAAGTGCCTCGCATGACCGCGGACCGCGGCTCTCTCGGATTTCCGCGTC 472
QY 141 SerGlyValArgValGlyProGluGlyAlaValAlaValTyrArgThrGluArgProGluVal 160
DB 473 TCCGCGTCCGAGTCTCGGCCCGCGCGCTGTGGCGGACGCGCGCGAGGTG 532
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAla 180
DB 533 AAGGCATGAGACACCGCGCGCTCCGCGCGTGTCTCCGCGTACGTGTCTATCGTTC 592
```

QY	180	aaspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyAspGlyValle	200
Db	593	CGACGAGGGATGGACTGCGGCTCGCGCGGGCGGTGTACGAGGAAAAGGTGGAGTCTT	652
QY	200	userGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe	220
Db	653	GTACACTGCTGCGGAGCTCAAGGCGTGGTTCGAGGAGCAGCCCTGTCTCCACTTCTACTC	712
QY	220	rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyG	240
Db	713	GGCGTCGATCTTCTGGCGTATGATGCTGCTGCAAGTGCAGAGCGAGGTGGGGTGG	772
QY	240	yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValleAspHi	260
Db	773	GGTAACAGTAGAGCTGGTGGACTTTGGCCATGTGCGCGAGGTGATGGGTGATTGACCA	832
QY	260	sAsnPheLeuGlyGlyLeuCysSerLeuLeuLeuPheValSerAspIleValProGluTh	280
Db	833	CAACTTCTTGGCGGGCTCTGCTAGCTGATCAAGTTTGTTTCTGACATTGTTCCAGAGAC	892
QY	280	rProHisThrGlnProLeuGlyProSer	289
Db	893	TCTTCAGAGCGACGCTTTGGGTCTCTCT	920
RESULT 4			
AX513564		1169 bp	DNA
LOCUS			linear
DEFINITION			PAT 05-OCT-2002
ACCESSION			Sequence 1 from Patent WO02059324.
VERSION			AX513564
KEYWORDS			AX513564.1 GI:23559664
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
US-10-042-894A-8 (1-289) x AX513564 (1-1169)			
QY	1	MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys	20
Db	84	ATGCCGACCTCCACCCCGGAGACCAAGTCGCGGCTCGGAGGCTCGGAGGCTCGGAGG	143



FEATURES  
source

Location/Qualifiers  
1..3416  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"  
72..407  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52959.1"  
/db\_xref="GI:23559685"  
/db\_xref="REMTREMBL:CAD52959"  
/translation="MPDLHPHQVAGHRASAKLGPLIDDSGLFYKPLQAGDGEHFE  
VAYFEARSAHAAPARIRDTIFFPFRHTRLUFTFAQGEPEPHLVLTDLLAGFEAPCV  
ADIKIGAIT"

CDS

ORIGIN

Alignment Scores:  
Pred. No.: 2 99e-89 Length: 3416  
Score: 1307.00 Matches: 260  
Percent Similarity: 90.34% Conservative: 2  
Best Local Similarity: 89.66% Mismatches: 9  
Query Match: 85.93% Indels: 20  
DB: 6 Gaps: 1

US-10-042-894A-8 (1-289) x AX513583 (1-3416)

QY	1	MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIys	20
Db	72	ATCCCGACCTCCACCCGCGGAGCACCAAGTCGCGGTCCACGCGCTCCGCCAGCAG	131
QY	21	LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrlsProLeuGlnAlaGlyAspArg	40
Db	132	CTGGCGCCACTCATCGACGACTCTGGCGCTCTTCTACAAGCGCTCCAGGCGCGGACCGT	191
QY	41	GlyGluHisGlnValAlaPheTyrlsGluAlaPheSerAlaHisAlaValProAlaArg	60
Db	192	GGGAGACACAGAGTCGCTTCTATGAGCGCGTCTCCGCCACGCGCGCTCCCGGCGCCG	251
QY	61	IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln	80
Db	252	ATCGAGACACTTCTTCCCGCGTTTCCAGGACAGGACCTCTCCGCCACGAGGCGCGAG	311
QY	81	ProGlyGluProHisProHisLeuValIleuAspAspLeuLeuAlaGlyPheGlnAlaPro	100
Db	312	CCCGGGAGCGCATCGCACCTCGTCTCTCGACGACCTCTCTCGCGGGGTTTGAGCGGCC	371
QY	101	CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr	120
Db	372	TGGTGTCCACACATCAAGATCGGTGCGCATCACGTG-----	406
QY	121	IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal	140
Db	407	-----ACACAGAGCGTTCTGTCTCGAATCCGGGTC	436
QY	141	SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal	160
Db	437	TCCGGCTCCGAGTCGTCCGCCCGCGAGGCGCGGTGTGGCGACGAGCGGCCCGAGGTG	496
QY	161	LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl	180
Db	497	AAGGCTATGACATTTGTCGGCGTCCGCCGCGGTGTCCGGCGCTACGTGTTCATCCGCTTC	556
QY	180	aSPbGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValIle	200
Db	557	CGACGAGGATGACTTCGCGCTTCGCGCGCGCGGTGTACGGAGAAAGGTGAGGTCCT	616
QY	200	userGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSer	220
Db	617	GTCAACGCTGCGGAGCTCAAGGCGTGTGTTCCAGGGGCGAGACTCTGTTCCACTTCTACTC	676
QY	220	rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyGly	240
Db	677	GGCGTCGATTTCTTCGGGCTATGATGCTGCTGAGTCTGCGACGCGGAGGTGGGGGTGG	736



Db 1163 GTGATTGACCAACTTCTTGGCGGGCTCTGCTCGCTCATCAAGTTTCATCGCGACATT 1104

Qy 277 ValProGlu 279  
||| |||

Db 1103 GTCGCCGAG 1095

RESULT 8  
AK072296

LOCUS AK072296 1570 bp mRNA linear PLN 24-JUL-2003

DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023018G11, full insert sequence.

ACCESSION AK072296

VERSION AK072296.1 GI:32982319

KEYWORDS FLJ\_CDNA; CAP trapper.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1

AUTHORS The Rice Full-Length cDNA Consortium, National Institute of  
Agricultural Sciences Rice Full-Length cDNA Project Team:  
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,  
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I.,  
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,  
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International  
Science Genome Sequencing & Analysis Group, Ootomo, Y., Murakami, K.,  
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,  
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,  
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,  
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,  
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:  
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Ootani, Y., Ota, Y.,  
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Yoshino, M., and Hayashizaki, Y.

TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from  
japonica rice

JOURNAL Science 301 (5631), 376-379 (2003)

MEDLINE 22752273

PUBMED 12869764

REFERENCE 2 (bases 1 to 1570)

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,  
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,  
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,  
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,  
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,  
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,  
Numasaki, R., Ohneda, E., Ohtsuki, K., Oka, M., Ooka, H.,  
Ootani, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,  
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,  
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,  
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,  
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,  
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and  
Yoshimura, A.

Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of  
Agricultural Sciences, Department of Molecular Genetics, Head of  
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki  
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,  
Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 28K full-length cDNA clones from japonica  
rice.  
URL : <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,  
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,  
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,  
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and  
Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y.,  
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,  
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center  
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,  
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,  
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,  
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ooka, H., Ootani, Y.,  
Oota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers  
1. 1570  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J023018G11"

ORIGIN

Alignment Scores:      Length:      1570  
Pred No.:      2,23e-77      Matches:      217  
Score:      1149.50      Conserv:      21  
Percent Similarity:      84.10%      Mismatches:      40  
Best Local Similarity:      76.68%      Indels:      5  
Query Match:      75.25%      Gaps:      3  
DB:      8

US-10-042-894A-8 (1-289) x AK072296 (1-1570)

Qy 2 SerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIysLeu 21

Db 384 TCCGACCTGGCGCGCGCGGAGCACAGTGGCGGGGCGCGCGCGCGGAGCGG 443

Qy 22 GlyProLeuIleAspGlySerGlyLeuPheTyrIysProLeuGlnAlaGlyAspArgGly 41

Db 444 GGCCCGCTGTCGACGGCGGAGGGGCTCTTCTACAGGCCCTTCCAGACGGGGGAGCGGG 503

Qy 42 GluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIle 61

Db 504 GAGCACGAGCGCGCTTCTACGGCGGCTTCACCGCGCACCGCGCGCTCCCGCGGTC 563

Qy 62 ArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnPro 81

Db 564 CGGCGCGCTTCTTCCCGCGGCTTCCAGCGCACCGCGCTTCTCCCGGGCGCCGAGCC 623

Qy 82 --GlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100

Db 624 GCGCGCGCGCTTACCCCGGACATCGTCTCGAGACCTCTCTCGGGGCTCCCGCTCCCC 683

Qy 101 CysValAlaAspIleLysIleGlyValaIleThrProProSerSerProGluProThr 120

Db 684 TGCCTCGCGAGCGTCAGATCGGCGCTTCGCGTGGCGCGCGCGATCCCGGACCCCTAC 743

Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140

Db 744 GTCGCCAAGTGCCTCGGCAAGGACCGGAGACCGGAGCGGCTCTCTCGGCTCCGCGTC 803

Qy 141 SerGlyValArgValValGlyProGlu--GlyAlaValThrArgThrGluArgProGlu 159

[illegible]

LOCUS	AX513580	643 bp	DNA	linear	PAT 05-OCT-2002
DEFINITION	Sequence 17 from Patent WO02059324.				
ACCESSION	AX513580				
VERSION	AX513580.1	GI:23559681			
KEYWORDS	Zea mays				
SOURCE	Zea mays				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE	Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.				
AUTHORS	Novel inositol polyphosphate kinase genes and uses thereof				
TITLE	Patent: WO 02059324-A 17 01-AUG-2002;				
JOURNAL	PIONEER HI-BRED INTERNATIONAL, INC. (US)				
FEATURES	Location/Qualifiers				
source	1. 643				
	/organism="Zea mays"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:4577"				

```

/ow_xref="taxon:457"

ORIGIN

Alignment Scores:
Pred. No.:          1.83e-52      Length:          643
Score:             808.50        Matches:         162
Percent Similarity: 93.14%       Conservative:    1
Best Local Similarity: 92.57%    Mismatches:     10
Query Match:       53.16%       Indels:         3
DB:                6           Gaps:           0

```

US-10-042-894A-8 (1-289) x AX513580 (1-643)

Qy	1	MetSer	Asp	Leu	His	Pro	Pro	Clu	His	Gln	Val	Ala	Gly	His	Arg	Ala	Ser	Ala	Ser	lys	20
Db	117	ATGCCCGAGCT	CCA	CCG	CCG	GAG	CAC	CAAT	CGCTCGGCT	CAC	CGCGCT	CTCCGCGCACAG	176								
Qy	21	Leu	Gly	Pro	Leu	Ile	Asp	Gly	Ser	Gly	Leu	Phe	Tyr	lys	Pro	Gln	Ala	Gly	Asp	Arg	40
Db	177	CTGGGCCCGCT	CAT	CGA	CGGT	CTGGGCT	CTT	CAC	AAG	CGGT	TCA	AGG	CGGCGCACCGT	236							

Qy	41	GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg	60
Db	237	GGGGAGCACGAGGTCGGCTTCTATGAGGCGCTTCTCGGCCACGCGNCGTCCGGCGCCGC	296
Qy	61	IleArgAspThrPhePheProArgPheHisGlyThrArgIleuLeuProThrGluAlaGln	80
Db	297	ATCCGAGACACCTTCTTCCCGCGTTCACGGCACGCGACTCTCTCCACCAGGAGCGCAG	356
Qy	81	ProGlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro	100
Db	357	CCGGGGAGCGGCATCCGACACTCTGTCGACGACCTCTCGCGGGTTGAGCGGCC	416
Qy	101	CysValAlaAspIleIysIleGlyAlaIleThrTyrProProSerSerProGluProTyr	120
Db	417	TGGTCGACAGACATCAAGATCGCGGCCCATCACGTGGCCACCGAGTTCGCGGAGCCCTAC	476
Qy	121	IleAlaIysCysLeuAlaIysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal	140
Db	477	ATGCNCAAGTACTTNGCCAAAGACCGCGGAGCCACGAGCGTTCGTGCGATTCCGGGTC	536
Qy	141	SerGlyValArgValValGlyProGluGlyAlaValIleTyrPargThrGluArg-ProGluVa	160
Db	537	TTGC--GTCCGAGTCGTGCGCCCGAGGGCGCGTGTGGCGGACGAGCGCCCGCGGGGT	594
Qy	160	LysAlaMetAspThr-AlaGlyValArgArgValLeuArg	173
Db	595	GAANGCTATGGACACCCGTCGNGNCGCGGNGTGTCTCGG	635

AX513576				PAT 05-OCT-2002
LOCUS	AX513576	1020 bp	DNA	linear
DEFINITION	Sequence 13 from Patent WO02059324.			

ORGANISM  
Parthenium argentatum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; campanulids; Asterales; Asteraceae; Asteroideae;  
Heliantheae; Parthenium.

1. Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.  
**AUTHORS**  
**TITLE** Novel inositol polyphosphate kinase genes and uses thereof  
**JOURNAL** Patent: WO 02059324-A 13 01-AUG-2002;  
**REFERENCE** PIONEER HI-BRED INTERNATIONAL, INC. (US)

```

CDS
11..908
/codon_start=1
/protein_id="CAD52957.1"
/db_xref="GI:23559678"
/ab_xref="EMBL:AF042957"
/note="unnamed protein product"

```

```

/db_ref="RSMTEMBL:CAD52957"
/translation="MLKADHQVAGHEAGLGLPLIDDSGRFYKPLQDNRGSESEA
FYESFSSNNNIEHIRKFPFYIGTAKMKASTDHPHVLQDLTSAHVNSVMDIKI
GRTWAPEASEAYIAKLCQDRBESTIPIGRISGLQVYIDDSGGYKPHRNVMRTQ
PADRVULLRKYVNSPSAEMMRITGLGDCSLASFYGGPNGILQAQMLKTFWEDT
IVHYFASCKFLVPEKRLVFGKARSNAEVKLIIDFAHYTDCNGVYIDHNFILGCLGSLIKFI
SDILSETSKDCNNGTNGOVR:"

```

## ORIGIN

Alignment Scores:	
Pred. No.:	3.02e-47
Score:	742.50
Percent Similarity:	67.83%
Best Local Similarity:	52.10%
Query Match:	48.82%
DB:	6
	5
	13
	79
	45
	149
	1020

US-10-042-894A-8 (1-289) x AX513576 (1-1020)

```
QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
Db 24 CTCAGGCCCCAGATCATCAGGTGCTGACATGAGCTGGGCTGGGAGCTTGCCCA 83
QY 24 LeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 84 CTCATTGATGATTCAGGCCGGTTTTACAAACCACTGCAGGGTGATAACCGTGGGTCAGAA 143
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgLileArgasp 63
Db 144 GAAGTAGCCCTTTATGAATCATTTCTTCTTAACTAATATATTCCAGAAACACATACGC--- 200
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 201 AATATCTTTCCTATATATTATGGCACCACCAATCATG-----AAGCATCCCTGGCTCT 254
QY 84 ProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103
Db 255 GACCATCTCTCATCGTGTGCAAGATCTTATCATCATGCTCATGTCAACCCATCTGTAATG 314
QY 104 AspLileLysIleGlyAlaLileThrTrpProProSerSerProGluProTyrLileAlaLys 123
Db 315 GACATCAAAATCGGCTCCAGACATCGGGCGCCAGAAAGCTTCCGAGCGGTACATTCGCAAA 374
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuPheArgValSerGlyVal 143
Db 375 TGTCTTAAAGAGATAGGAAAGACCAAGTATTCATTGGGATTGAGGATTCGGGGCTG 434
QY 144 ArgValValGlyProGluGlyAla---ValTrpArgThrGluArgProGluValLysAla 162
Db 435 CAAGTCTATATCATGATCGGTCAGGGTTTTATAGCCCTCATAGAAATTTACATGCGTAA 494
QY 163 MetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAlaAspGlu 182
Db 495 ACCGGCCAGCTGATGTAGACTACTTCTTAGGAAATTTGTTCTTCTTAACCGCTGCA 554
QY 183 -----GlyMetAspCysAlaLeuAlaAlaValTyrGly 194
Db 555 GAGATGGAAATCGGCACAGGCCCTAGCCCGGATGTCTTTAGCATCTTTGTTATGT 614
QY 195 GlyLysGlyValLysSerGlnLeuArgGluLeuLysAlaTyrPheGluGlnThr 214
Db 615 GGGCCTAATGGGATATTAGTCAACTGATGGAATTTGAACATGCTTTGAAGATCAACA 674
QY 215 LeuPheHisPheTyrSerAlaSerLileLeuGlyTyrAspAlaAlaValAlaAla 234
Db 675 ATTTACCACTCTATGCTGTGTTCTTTTGTTCATCTTTGAAAGAGGGTTGGTGTAA 734
QY 235 GlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGly 254
Db 735 GGTGCT---CGGTCAACCGCAGAACTCAATATTGATTGCTCATGTACATGATG 791
QY 255 AspGlyValIleAspHisAlsnPheLeuGlyGlyLeuCysSerLeuLileLysPheValSer 274
Db 792 AATGTGTTATTGATCAACAATTTCTTGGTGGGCTCTGTCTTGTATGATAAGTTTCATCT 851
QY 275 AspLileValProGluThr 280
Db 852 GACATACCTTTCGGAGACA 869
```

## RESULT 11

```
AX513574
LOCUS AX513574
DEFINITION Sequence 11 from Patent WO02059324.
ACCESSION AX513574
VERSION AX513574.1 GI:235559674
KEYWORDS Eucalyptus grandis
SOURCE Eucalyptus grandis
ORGANISM Eucalyptus grandis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
```

rosids; Myrtales; Myrtaceae; Eucalyptus.

## REFERENCE

1 Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.  
Novel inositol polyphosphate kinase genes and uses thereof  
Patent: WO 02059324-A 11 01-AUG-2002;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)

## FEATURES

## source

1. 1195  
/organism="Eucalyptus grandis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:71139"  
116. 1048  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52956.1"  
/db\_xref="GI:23559675"

## CDS

/translation="MLKVPDQVAGHRDGGKGLPVDSDGRFYKPLQSDHRRGDTVEA  
FYEFYSNTEIPGHIRKEFPFAPFGTKTTEASDGGPQPHLVLEIDVSGRTNPSLMDIK  
TGSRTWPEASEEVIQKLEKDRNSTVSLGFRISGLRVYQNSAAGFWQPKKVVYSF  
NADGVSALRKRSESNLSLGNVDPDCLYASKYCHRGGLAQLQLKEWFEVQNYH  
FYSCSLIILYDRSALDGCARPKVKLVDFAHVMDGHGVIDHNFLGLGSLCVIRKIDIA  
DEDNKAACEVNLGLKENGIFYKSTPELDHEAC"

## ORIGIN

## Alignment Scores:

Pred. No.:	1.18e-45	Length:	1195
Score:	722.50	Matches:	149
Percent Similarity:	66.31%	Conservative:	38
Best local Similarity:	52.84%	Mismatches:	84
Query Match:	47.50%	Indels:	11
DB:	6	Gaps:	6

US-10-042-894A-8 (1-289) x AX513574 (1-1195)

```
QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
Db 119 CTCAGGTCCTGGATCATCAAGTCGCGGTCCACGGGGAGAGCGGGGAAAGCTGGGCCA 178
QY 24 LeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 179 CTGGTGGATGATTCGGGCGCTTCTATAGGCTCTCCAGAGCGATCATCGCGAGACAG 238
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgLileArgasp 63
Db 239 GAAGTGGCTTTTACGAGTCAATCTATTCCAAATCCGAGATCCCGAGTCCACATTCGC--- 295
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 296 AAATCTTCTCTGCTTTCACGGAACCTAAGACTATT-----GAGGCGTCTGATGATCG 349
QY 84 ---ProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysVal 102
Db 350 GGTCTCAACCTCACTGCTGTTCTGGAGGATCTGCTCGGGTCCGACGAAACCATCTCTC 409
QY 103 AlaAspLileLysIleGlyAlaLileThrTrpProProSerSerProGluProTyrLileAla 122
Db 410 ATGNCATCAAGACTGGATCCAGAACATGGTATCCGAGGCGCTCTGAGAGGTACATCA 469
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
Db 470 AAGTGTCTTAGAGAAAGATCGAAATAGCACAACGCTTTCATTTGGGTTTTCAGGATTTCTGG 529
QY 143 ValArgValVal---GlyProGluGlyAlaValTrpArgThrGluArgProGluValLys 161
Db 530 CTAAGGGTATATCAAAATAGCAAGCTGGATTTTGGCAACCTGAGAAAGAGGTTGTTAT 589
QY 162 AlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAlaAsp 181
Db 590 AGCTTTAATGCGGCGGTGTGAGTCCGCTCTGAGGAAGTTGTTCTTCTTCAACTGTCT 649
QY 182 GluGly-----MetAspCysAlaLeuAlaAlaValTyrGlyLysGly 197
```





CDS  
PSLAETKLHLKLVIGQCNIIGIYGLANLGSLSHLISGCPCLTVVBSLRLNLGTL  
LELSGYTITNLPSSLTYTKLSKLVSDSOLPDLTNLKNLRCIKLCSGDNFIETGL  
HTLSLELRVWGSIRKLDLTGLVKLEILAFDSTQLTETRGGLSGLRLHWSRC  
QSIKELPNLSGLKLSYILLEKCRHLKEVYRLEELKRLUDFNTHRG"  
17600..18457  
/function="unknown"  
/note="ORF  
related to human inositol hexakisphosphate kinase 2  
protein"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAC35324.1"  
/db\_xref="GI:13509210"  
/db\_xref="GOA:Q9AR40"  
/db\_xref="SPTREMBL:Q9AR40"  
/translation="MLKVPERQVAGHQDINGLLGLVDDSGRFYKPLQCDREGATEMA  
FYTSSTNKTIPQVRRFFQAFQGSIEASDGGSLHPLHLEDLTSLRLHLCVMDIK  
IGSRWYTPASQAYIEKCKXDVSSNELLFRISGLQVNGKESSEVWLKPKLL  
LNLNADEVLRVLRVSNLXSDPDSCFAAVSVSGSLGILQALLEKAWFEDQTIYH  
FNSCSVLMLEYEKETKMNVGNGEESLGGAAVLIDFAHVTEGNGVIDHFNGLGLSLIK  
FISENQIML"  
join(20265..20815,21104..22202,22415..22690,22818..24218)  
/gene="Ngc-B"  
join(20265..20815,21104..22202,22415..22690,22818..24218)  
/gene="Ngc-B"  
/note="N rust resistance gene homolog: TIR-NBS-LRR  
protein"  
/codon\_start=1  
/product="Ngc-B protein"  
/protein\_id="CAC35325.1"  
/db\_xref="GI:13509211"  
/db\_xref="GOA:Q9AR40"  
/db\_xref="SPTREMBL:Q9AR40"  
/translation="MRSDDSSISGSLRSVSDPTLPLPSGEYEVFLSPRGPVDRQ  
TFADLYXAWLVRSKIRTPDEEGKQGETIGSSLIQAATESKIYIPILITQNYASSKWC  
LOELAKWVDKNGGGGKGQHIILPVFMDPRDVRHPSDGPYKEAFQHNLDKHPET  
ILEWKEALQDVYKMGWHINELTQGAVDVKIFTTIEPHLRANTLATDELVGDSSV  
EVMELMNLDSHSTSEIRIIGYMGGLGKTLAKAVENKVSQPERCCFLDNI RETLAR  
NDGVVALONKVIIDILRKSDQAKNASDGVILIBERVRRHKIVFLVDIDIESFHEDEI  
FGKLGDSFTRSLPITRDARTLELNECKMFGLEMSHDSLQFSGHAPGVDPPE  
DYASLCRFIVAGSLPALKVIGSLIFKSKRFWEDEKLIELKAIKSAKVQERLKVSY  
NLTHNEKQIFLDIACLVGAKGVPMYMSDCLYIPASTLRTLVQSLRMDNDKRF  
WMHDIRDLGRTIYRENSQNPYKRSIWSNNDAILKNGENDGVLEALRVMKMGEG  
YALTNKQFQSRFLVLENDLSGNFKNILPNLRVLRVYRGPSPSGNLNPKLVIL  
ELDCVYTHSGWNEIKAKGLKVNILTSOGILEKVPDLSTCGELCLFLPHKQWR  
GELDTGFKDLKVLIDNQTEITILKGEVESLQNLQDLVGRSGLIEVPAGISKUSLE  
FLDLSVAKHDEVMPLNGKLVISFSLSALPSSLIKIDICDSRNLRQRPNLASVTN  
LTRHLKVEHLIEPGLGKLLBSLSICNAPNLDNLDGLNVLKLELALERCPILG  
KLPSAELTKHLKVRWCVDVIGIYGLNGLDLSHLIDISWCPRLTVMOLLHSLILK  
GTVSSGFELTNILPLSLISYTKLRTLEVRSSQLPDLTNLKNLRLDITITGCRELIETL  
GLHTLESLELSMERCPVSKLDLGLIKLTIHICTOLTETRGGLGLESLOLTFM  
SGCSQIKELPNLSGLKLSYILLEKCRHLKEVYRLEELKRLUDFNTHRG"  
SRKQQLVTSQARALFINN"  
ORIGIN  
Alignment Scores:  
Pred. No.: 8,32e-44 Length: 25054  
Score: 718.50 Matches: 149  
Percent Similarity: 66.78% Conservative: 44  
Best local Similarity: 51.56% Mismatches: 83  
Query Match: 47.24% Indels: 13  
DB: 8 Gaps: 7  
US-10-042-894A-8 (1-289) x LUS310150 (1-25054)  
Qy 7 ProGluHisGlnValAlaGlyHisArgAlaSerLysLeuGlyProLeuLeuAsp 26  
Db 9625 CCGGAACATCAGGTGCTGGACATCAAGCCATCAAGCGCTCTCTCGGTCTCTTGTAGAC 9684  
Qy 27 GlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHisGluValAla 46  
Db 9685 GATTGAGGGCGATTCTACAGCCACTTCAGGCGCATGAACGTGGGCAACCGAGATGCA 9744

Qy 47 PheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArgAspThrPhePhe 66  
Db 9745 TTCACACATCTTTTCGTCTCAAAAAGATTCACAAACAGTCGCGG---AGGTTTTC 9801  
Qy 67 ProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGluPro---His 85  
Db 9802 CCGGCTTCATCGCACTCAAAAGTATA-----GAGGCTTCTGATGGATCTGGCCTCCAC 9855  
Qy 86 ProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysValAlaAspIle 105  
Db 9856 CCGCACCTCATCTTAGAAGATCTCACGTCGATCGCTGGTTGCATCATGTGTATGACATC 9915  
Qy 106 LysIleGlyAlaIleThrTipProSerSerProGluProTyrIleAlaLysCysLeu 125  
Db 9916 AAGATTGGTTCAGGACATGATGATCCCGAGGCTTCTCAAGCTTACATCGAAGATGCTG 9975  
Qy 126 AlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgVal 145  
Db 9976 AAGAAGATGTGGAATCAAGCAGTCCCTTCTCGGGTTTAGGATATCCGGATTGCGAGTT 10035  
Qy 146 ValGlyProGlu-----GlyAlaValTrpArgThrGluArgProGluValLys 161  
Db 10036 TACGGTAACGCAAAAGAAATCATCTGAGGTGCTCTTGAAGCCTGAGAGGAGCTTCTCCAG 10095  
Qy 162 AlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer---ValAla 180  
Db 10096 AATCTTACTGCTGATGAAGTTAGCTGCTGTGCTTAAACGGTTTGTCTTCCACCCGAG 10155  
Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLeu 200  
Db 10156 TCGGACCAACCGGATTGTTGCTTTCGAGCGGTGCTTTCAGGTGCTTTCGATGGGATCTG 10215  
Qy 201 SerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSer 220  
Db 10216 GCTCAGTTGCTGGAGCTGAAGCGGTGCTTTCAGGTGCTTTCGATGGGATCTTCACTCT 10275  
Qy 221 AlaSerIleLeuLeuGlyTyr-----AspAlaAlaAlaValAlaAlaGlyValAspGly 238  
Db 10276 TGTTCTGTTCTTATGCTGTATGAGAGGAGAGGAGAGGAGTGTGATGGAGGAGAGAG 10335  
Qy 239 -----GlyGlyValThrValLysLeuValAspPheAlaHisAlaGluGlyAspGly 256  
Db 10336 AGTTGGGTTGCTGCTGCGGTGAAGCTCATCGATTTCATGTCGAGGAGGAGGAGGAG 10395  
Qy 257 ValIleAspHisAnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIle 276  
Db 10396 GTTATCATCACAACTTCTTGGAGGCTCTGCTCATTTGATATAAGTTTCAATTCGAGATC 10455  
Qy 277 ValProGluThrProHisThrGlnPro 285  
Db 10456 CTCACCGGTCGGATGATGAATTCCTCCG 10482  
RESULT 13  
ATH243592 903 bp DNA linear PLN 01-JUL-2001  
LOCUS Arabidopsis thaliana gene for inositol-(1,4,5) trisphosphate  
DEFINITION 3-kinase.  
ACCESSION AJ243592  
VERSION AJ243592.1 GI:14588984  
KEYWORDS inositol-(1,4,5) trisphosphate 3-kinase.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
1  
Xia,H.J., Breatley,C. and Mueller-Roeber,B.  
REFERENCE Identification of inositol-(1,4,5) trisphosphate 3-kinase from  
AUTHORS Arabidopsis thaliana  
TITLE Arabidopsis thaliana  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 903)  
AUTHORS Xia,H.J.  
TITLE Direct Submission



JOURNAL Submitted (09-JUL-1999) Xia H.J., AG Mueller-Roeber,  
Max-Planck-Institute of Molecular Plant Physiology,  
Karl-Liebknecht-Strasse 25, Haus 20, 14476 Golm, GERMANY

FEATURES Location/Qualifiers

source 1..903

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/variety="Col"

/db\_xref="taxon:3702"

CDS 1..903

/function="phosphorylation of the second messenger  
inositol-(1,4,5) trisphosphate"

/codon\_start=1

/product="inositol-(1,4,5) trisphosphate 3-kinase"

/protein\_id="CAC43070.1"

/db\_xref="GI:14588985"

/db\_xref="GOA:Q9FLT2"

/db\_xref="SPTREMBL:Q9FLT2"

/translation="MLKVPBHVAGHISADGKGLPIVDQGRFFKPLQDGRSGHEAK  
FYSEFTSNMKVPDHIHRYFVYVHTQGLVEASDGSGLPHLDVLDVVGVPANPSVMDVK  
IGRTWYLPDYSEYFKKCIKKDKTTTSLGSPRVSGFKIPDHQESSFWRAEKKVLGY  
NADGAKPLDKFKFVSSNSPDSNLTNCPAFSEVYGCNGILAOILLKQWFEQTQTLVY  
FNCSILMIYENSIILMOGDDAPAPRAQVKLVDPFHVLDNGVGDHNFGLGLCSFIK  
FIKDLIQSVEKHDETTSTLENR"

ORIGIN

Alignment Scores:

Pred. No.:	3,41e-45	Length:	903
Score:	714.50	Matches:	150
Percent Similarity:	66.55%	Conservative:	39
Best Local Similarity:	52.82%	Mismatches:	82
Query Match:	46.98%	Indels:	13
DB:	8	Gaps:	8

US-10-042-894A-8 (1-289) x AT4243592 (1-903)

QY	4	LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro	23
Db	4	CTCAAGTCCTCAACACCAAGTTGCTGGTCACATTGCTAGTATGGAAGCTCGGTCCA	63
QY	24	LeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis	43
Db	64	CTCGTAGATGACCAAGCCGGTCTTCAAGCCACTTCAGGAGATTCGTGGCGAACAC	123
QY	44	GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgileArgAsp	63
Db	124	GAGCTTAAGTTCTATGAGTCTTTCACATCGAATGAGGTTCACATCATCATCAT	180
QY	64	ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu	83
Db	181	AGATACCTCCGGTGTATCAGCGCACTCAGCTAGTT-----GAAGCATCTGATGATCT	234
QY	84	ProHis---ProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysVal	102
Db	235	GGCAAGTTCCTCATCTGTCTTGATGATGTGTTCAGGATGCGAACCCTCGGTA	294
QY	103	AlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAla	122
Db	295	ATGGATGTTAAGATTGGATCTAGACATGGTACCCGGATGTATCAGAAGAACTACTCAAG	354
QY	123	LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly	142
Db	355	AAATGTATTAAAGAAATAGACAGACACCACCGGTTTCGTGGGGTTTCAGGGTTTCAGGT	414
QY	143	ValArgValValGly---ProGluGlyAlaValTrpArgThrGluArgProGluValLys	161
Db	415	TTTAAATTTTGGATCACCAGGAATCAAGTTTTCAGGAGCTGAGAGAAGCTTGTGTTCT	474
QY	162	AlaMetAspThrAlaGlyValArgGValLeuArgTyrValSer-----SerVal	179
Db	475	GGGTATATGCATGGTGTAGATGGCTCTGAGGAAGTTTGTGTATCGAACTCTCC	534
QY	180	AlaAspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGly	197

Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

#### FEATURES

source

Location/Qualifiers  
1..903  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="U13510"  
/ecotype="Columbia"  
/note="This clone is in pUNI 51"  
1..903  
/notes="unknown protein"  
/codon\_start=1  
/product="AT5G61760/mac9\_50"  
/protein\_id="AAL62012.1"  
/db\_xref="GI:18252263"  
/translation="MLKVPHEQVAGHIASDGKLGILPVDQGRFFKPLQDSRGEHEAK  
FYESTSNMKVPDHIHRYFPVYHGTQVLEASDGSKLPHLVLDVVGYPNSVMDVK  
IGRTWYPOVSSEYFKCIKKDQRTTTSVLSGRVSGFKI FDQESSFWRAEKLVLGY  
NADGARLARKFVSSNSPADSNLTPNCAPASEVYGGCGILAQLELKDQWFTQTLVH  
FNSCSILMIYENESILMQGGDAPAPRAQVKLVDFAHVLIDGNGVIDHNFLLGLCSFIK  
FIKDILQSVKXHDFTDTSLENGR"

CDS

1..903  
/notes="unknown protein"  
/codon\_start=1  
/product="AT5G61760/mac9\_50"  
/protein\_id="AAL62012.1"  
/db\_xref="GI:18252263"  
/translation="MLKVPHEQVAGHIASDGKLGILPVDQGRFFKPLQDSRGEHEAK  
FYESTSNMKVPDHIHRYFPVYHGTQVLEASDGSKLPHLVLDVVGYPNSVMDVK  
IGRTWYPOVSSEYFKCIKKDQRTTTSVLSGRVSGFKI FDQESSFWRAEKLVLGY  
NADGARLARKFVSSNSPADSNLTPNCAPASEVYGGCGILAQLELKDQWFTQTLVH  
FNSCSILMIYENESILMQGGDAPAPRAQVKLVDFAHVLIDGNGVIDHNFLLGLCSFIK  
FIKDILQSVKXHDFTDTSLENGR"

#### ORIGIN

Alignment Scores:

Pred. No.: 3 41e-45 Length: 903  
Score: 714.50 Matches: 150  
Percent Similarity: 66.55% Conservative: 39  
Best Local Similarity: 52.82% Mismatches: 82  
Query Match: 46.98% Indels: 13  
DB: 8 Gaps: 8

US-10-042-894A-8 (1-289) x AY072621 (1-903)

QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerLeuGlyPro 23  
DB 4 CTCAGGTCCTCGAACCAAGTGTGCTGCATCTGCTAGTGGGAGCTCGGTCA 63  
QY 24 LeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43  
DB 64 CTCGTAGATGACCAAGCGCGGTCTCTCAAGCCACTTCAGGAGATCTCGTGGCGAACAC 123  
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63  
DB 124 GAGGTAAGTCTTGTAGTCTTTCATCGAACATGAGGTTCCAGATCACATCCAT--- 180  
QY 64 ThrPheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83  
DB 181 AGATACTTCCCGGTGTATCAGCGCACTCAGTAGT-----GAGCATCTCATGATCT 234  
QY 84 ProHis---ProHisLeuValLeuAspAspLeuLeuAlaGlyPheClnAlaProCysVal 102  
DB 235 GCAAGACTTCTCTCATCTGTTCTTGTATGATGTTGTTTTCAGGTTACGCAACCCGCGTA 294  
QY 103 AlaAspIleLeuValGlyAlaIleThrTrpProSerSerProGluProTyrIleAla 122  
DB 295 ATGGATGTAAAGATTGGATCTAGGACATGTTACCGGATGATACAGAAATCTTCAG 354  
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142  
DB 355 AAATGATTAAAGAAATAGACAGACCACCGGTTTCGTTGGGTTTCAGGTTTCAGGT 414  
QY 143 ValArgValValGly---ProGluGlyAlaValTrpArgThrGluArgProGluValLys 161  
DB 415 TTTAAGATTTTTCATCACCAGAAATCAAGTTTTTGGAGAGCTGAGAAAGCTTGTCTT 474  
QY 162 AlaMetAspThrAlaGlyValArgValLeuArgValTyrValSer-----SerVal 179  
DB 475 GGGTATATGCAGATGGTGTAGATTGGCTCTGAGGAAGATTGTGTCTCGAACTCTCCC 534

QY 180 AlaAspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGly 197  
DB 535 GCTGATCTTAACCTTGACACCAACTGTCTTTTGCATCAGAGGTTTATGCGGTGTGAAC 594  
QY 198 GlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHis 217  
DB 595 GGGATCTTAGCGCAGTTGTTGGAGCTTAAAGATTGGTTCGAAACCCCAACGCTTACCAT 654  
QY 218 PheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaVal---AlaAlaGlyGly 236  
DB 655 TTTCAATTCCTGCTCGATCTGATTTACGAGATCAATCAATCTTGATGCAAGGAGGA 714  
QY 237 AspGlyGly-----GlyValThrValLysLeuValAspPheAlaHisValAlaGlu 253  
DB 715 GATGATGACCGGCACCGGCACAGTGAAGCTGTTGATTTCCGGCATGTTCTTGAT 774  
QY 254 GlyAspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheVal 273  
DB 775 GGAACCGGTGTCATCGACCATTAATTTCTTGGGTGGACTCTGCTCTTCATAAAGTTTCATC 834  
QY 274 SerAspIleVal 277  
DB 835 AAAGATATTTCTT 846

#### RESULT 15

AY147936

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

note="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="5"

/ecotype="Columbia"

1..903

/gene="IPK2b"

1..903

/note="inositol phosphate kinase"

/codon\_start=1

/product="inositol polyphosphate 6-/3-/5-kinase 2b"

/protein\_id="AAN63056.1"

/db\_xref="GI:24850181"

/translation="MLKVPHEQVAGHIASDGKLGILPVDQGRFFKPLQDSRGEHEAK  
FYESTSNMKVPDHIHRYFPVYHGTQVLEASDGSKLPHLVLDVVGYPNSVMDVK  
IGRTWYPOVSSEYFKCIKKDQRTTTSVLSGRVSGFKI FDQESSFWRAEKLVLGY  
NADGARLARKFVSSNSPADSNLTPNCAPASEVYGGCGILAQLELKDQWFTQTLVH  
FNSCSILMIYENESILMQGGDAPAPRAQVKLVDFAHVLIDGNGVIDHNFLLGLCSFIK  
FIKDILQSVKXHDFTDTSLENGR"

ORIGIN

Job time : 4547 secs

## Alignment Scores:

Pred. No.:	3,41e-45	Length:	903
Score:	714.50	Matches:	150
Percent Similarity:	66.55%	Conservative:	39
Best Local Similarity:	52.82%	Mismatches:	82
Query Match:	46.98%	Indels:	13
DB:	8	Gaps:	8

US-10-042-894A-8 (1-289) x AY147936 (1-903)

QY	4	LeuHisProGlnHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro	23
DB	4	CTCAAGTCCTCGAACACCAAGTGTCTGTCACATTGCTAGTATGGGAAGCTCGGTCCA	63
QY	24	LeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis	43
DB	64	CTCGTAGATGACCAAGCGCGTCTTCAAGCCACTTCAGGGAGATTCTCGTGGCGAACAC	123
QY	44	GluValAlaPheTyrGlnAlaPheSerAlaHisAlaValProAlaArgIleArgAsp	63
DB	124	GAGGCTAAGTCTATGATGCTTTCACATCGACATGAAGGTTCCAGATCACATCCAT---	180
QY	64	ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGlnAlaGlnProGlyGlu	83
DB	181	AGATACCTTCCCGTGTATGATACCGGACCTCAGCTAGTT-----GAAGCATCTGATGGATCT	234
QY	84	ProHis---ProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaProCysVal	102
DB	235	GGCAAGCTTCTCATCTTGTCTTGTGATGATGTTGTTTCAAGGTACGCAACCCGTCGGTA	294
QY	103	AlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAla	122
DB	295	ATGGATGTTAAGATTGGATCTAGGACATGATCGCGATGTATCAGAAGATACTTCAAG	354
QY	123	LysCysLeuAlaLysAspArgGlyThrSerValLeuLeuGlyPheArgValSerGly	142
DB	355	AAATGATTAAAGAAAGACAGACACACCGCTTTCGTTGGGGTTTCAGGGTTTCAGGT	414
QY	143	ValArgValValGly---ProGluGlyAlaValTyrArgThrGluArgProGluValLys	161
DB	415	TTTAAGATTTTTCATCACCAGATCAAGTTTGGAGGCTGAGGAAGCTTGTCTT	474
QY	162	AlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSer-----SerVal	179
DB	475	GGGTATAATGCAGATGGTGTAGATTGGCTCTGAGGAAGTTTGTGTATCGAACTCTCCC	534
QY	180	AlaAspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyLysGly	197
DB	535	GCTGACTCTAATGTGACACCAACTGTCTTTGTCATCAGAGTTTATGGCGTTGTAAC	594
QY	198	GlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHis	217
DB	595	GGGATCTTAGCGCAGTTGTTGGAGCTTAAAGATTGTTGMAACCCAAACGCTTTACCAT	654
QY	218	PheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaVal---AlaAlaGlyGly	236
DB	655	TTCAATTCTGCTCGATTCTGATGATTACGAGATGAATCAATCTTGATGCAAGGAGGA	714
QY	237	AspGlyGly-----GlyValThrValLysLeuValAspPheAlaHisValAlaGlu	253
DB	715	GATGATGCCCGGCACCGGACCAAGTGAAGCTGGTGAATTCGCGCATGTTCTTTGAT	774
QY	254	GlyAspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheVal	273
DB	775	GGAAACGGTGTATCGACCATATTTCTGGGTGGACTCTGCTCTTTCATAAAGTTTCATC	834
QY	274	SerAspIleVal	277
DB	835	AAAGATATTCTT	846

Search completed: March 27, 2004, 07:35:02

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 20:51:49 ; Search time 18 Seconds  
(without alignment)  
836.015 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVAGHRASAK.....IKFVSIVPETHPTQLGPS 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145	9.5	461	1	IP3K HUMAN
2	131	8.6	459	1	IP3K RAT
3	123	8.1	355	1	ARG3 YEAST
4	114.5	7.5	872	1	IP3L RAT
5	109.5	7.2	946	1	IP3L HUMAN
6	97.5	6.4	362	1	ADD_MYCLE
7	93.5	6.1	372	1	CARA_SULTO
8	89.5	5.9	331	1	NAGZ_XANCP
9	89	5.9	380	1	GRK_RACHD
10	88.5	5.8	352	1	EGSA_THEAC
11	88.5	5.8	370	1	DNRJ_STRPE
12	88.5	5.8	377	1	VTBX_AGRF5
13	87.5	5.8	510	1	ER53_CERAE
14	87.5	5.8	550	1	CHIT_NPVOP
15	86.5	5.7	373	1	CSZ2_MYCTU
16	86.5	5.7	377	1	VTBX_AGRF9
17	86.5	5.7	454	1	UCR2_NEUCR
18	86.5	5.7	1032	1	YGFK_ECOLI
19	86	5.7	364	1	X880_MYCTU
20	85.5	5.6	334	1	NAGZ_XANCP
21	84.5	5.6	447	1	DCDA_MYCTU
22	84	5.5	316	1	PECT_ERWCH
23	84	5.5	411	1	DHE3_ARATH
24	83.5	5.5	304	1	PQOB_PSEAE
25	83.5	5.5	1201	1	MED_WYXXA
26	83	5.5	131	1	PRO3_LILLO
27	83	5.5	492	1	DP22_HUMAN
28	82.5	5.4	485	1	PRLY_HA1N1
29	82	5.4	131	1	AR22_ARATH
30	82	5.4	310	1	MTAA_ERUSU
31	82	5.4	541	1	G6P1_DEIRA
32	82	5.4	742	1	UL47_HSVBP
33	81.5	5.4	529	1	IMDH_MYCTU

34 81.5 5.4 884 1 RPOB\_NPVOP O12934 oryza pseu  
35 81 5.3 411 1 DHE1\_ARATH Q43314 arabidopsis  
36 80.5 5.3 298 1 KHSE\_ANASP Q8VZV9 anabaena sp  
37 80.5 5.3 352 1 ADH\_THERB P14941 thermoanaer  
38 80.5 5.3 354 1 TRD1\_STRCO O68608 streptomyc  
39 80.5 5.3 391 1 Y532\_METJA Q57952 methanococ  
40 80.5 5.3 1364 1 PUR2\_DROPS P16340 d trifuncti  
41 80 5.3 294 1 PANE\_ARCFU O28578 archaeoglob  
42 80 5.3 311 1 MTAA\_MYCLE P46811 mycobacteri  
43 80 5.3 411 1 DHE2\_ARATH Q38946 arabidopsis  
44 80 5.3 411 1 DHE3\_VITVI P52596 vitis vinif  
45 80 5.3 421 1 FABF\_RHIME P56902 rhizobium m

#### ALIGNMENTS

RESULT 1  
IP3K HUMAN  
ID IP3K HUMAN STANDARD; PRT; 461 AA.  
AC P23677;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Inositol-trisphosphate 3-kinase A (EC 2.7.1.127) (Inositol 1,4,5-  
trisphosphate 3-kinase) (IP3K) (IP3 3-kinase).  
GN ITPKA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=91128380; PubMed=1847047;  
RA Takazawa K., Perret J., Dumont J.E., Erneux C.;  
RT "Molecular cloning and expression of a human brain inositol 1,4,5-  
trisphosphate 3-kinase.";  
RL Biochem. Biophys. Res. Commun. 174:529-535(1991).  
[2]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Brain;  
RX MEDLINE=91088302; PubMed=2175886;  
RA Takazawa K., Perret J., Dumont J.E., Erneux C.;  
RT "Human brain inositol 1,4,5-trisphosphate 3-kinase cDNA sequence.";  
RL Nucleic Acids Res. 18:7141-7141(1990).  
CC -!- CATALYTIC ACTIVITY: ATP + 1D-myo-inositol 1,4,5-trisphosphate =  
ADP + 1D-myo-inositol 1,3,4,5-tetrakisphosphate.  
CC -!- ENZYME REGULATION: IP3K is activated by calmodulin.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X54938; CAA38700.1; --  
DR PIR; JN0129; JN0129.  
DR Genew; HGNC:6178; ITPKA.  
DR MIM; 147521; --  
DR GO; GO:0008440; P:inositol-trisphosphate 3-kinase activity; TAS.  
DR GO; GO:0007165; P:signal transduction; TAS.  
DR InterPro; IPR005522; IPK.  
DR Pfam; PF03770; IPK; 1.  
DR Transferase; Kinase; Calmodulin-binding.  
FT DOMAIN 287 295 CALMODULIN-BINDING (BY SIMILARITY).  
SQ SEQUENCE 461 AA; 51008 MW; 18CA214A091F5B19 CRC64;

Query Match 9.5%; Score 145; DB 1; Length 461;  
Best Local Similarity 23.5%; Pred. No. 0.0001;  
Matches 76; Conservative 40; Mismatches 97; Indels 110; Gaps 15;

```
QY 10 QVAGHRASAKGLPLIDSGSLFYKPLQAGDRGEHEVAFYAFSAHAAPVARIIDTFFPRF 69
Db 190 QLAGHTGSGFAAGT-----SGLLKRCSEPER--YCLARLMADALRGCVPA-----F 234
QY 70 HGTRELLPTEAQPGEHPHLVLDLLAGFQAPCVADIKIGAITW----- 112
Db 235 HG-----VVERDGB--SYLQQLDLDGDFGCVLDCKMGVRYLYEELTKARERPKLRKD 287
QY 113 -----PPSP--EPYIAKCLAKDRG-----TTSVLLGFRVSGVRVV--GPEGAVWR 154
Db 288 MYKMLAVDEAPTEEHEAQRVTKRYMQWRREGISSSTTLGFRIGIKKADGSCSTDFK 347
QY 155 TERPEVKAMD-----TAGVRVLRVRYSSVADEGMDCALAAAVYGGKGVLSOLRELKAW 209
Db 348 TTRSREQVTRVFEFVQGDDEVLRV-----LNRLQOIRDT 383
QY 210 FEEQTLFHFY---SASILLGVDAAAVAGDGGGVTVKLVDAEHAFAV--GDG-VIDH--- 260
Db 384 LEVSEFFRRHEVIGSSLLFVHDCHFRAG-----VWLIDFGKTTPLPDGQILDHRRPW 435
QY 261 -----NFLGGCLSLIKFVSDI 276
Db 436 EEGNREDGYLLGLDNLIGILASL 458

RESULT 2
IP3K_RAT
ID IP3K_RAT STANDARD; PRT; 459 AA.
AC P17105;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-1-trisphosphate 3-kinase A (EC 2.7.1.127) (Inositol 1,4,5-
triphosphate 3-kinase) (IP3K) (IP3 3-kinase).
GN ITPKA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=91090700; PubMed=21176078;
RA Takazawa K., Vandekerckhove J., Dumont J.B., Erneux C.;
RT "Cloning and expression in Escherichia coli of a rat brain cDNA
RT encoding a Ca2+/calmodulin-sensitive inositol 1,4,5-trisphosphate
RT 3-kinase."
RL Biochem. J. 272:1107-112 (1990).
[2]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=90208336; PubMed=2157285;
RA Choi K.Y., Kim H.K., Lee S.Y., Moon K.H., Sim S.S., Kim J.W.,
RA Chung H.K., Rhee S.G.;
RT "Molecular cloning and expression of a complementary DNA for inositol
RT 1,4,5-trisphosphate 3-kinase."
RL Science 248:64-66 (1990).
CC -!- CATALYTIC ACTIVITY: ATP + 1D-myo-inositol 1,4,5-trisphosphate =
CC ADP + 1D-myo-inositol 1,3,4,5-tetrakisphosphate.
CC -!- ENZYME REGULATION: IP3K is activated by calmodulin.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR Pfam; PF03770; IPK; 1.
KW Transferase; Kinase; Calmodulin-binding.
FT DOMAIN 285 293 CALMODULIN-BINDING.
SQ SEQUENCE 459 AA; 50870 MW; 8093DADC0FADC290 CRC64;

Query Match 8.6%; Score 131; DB 1; Length 459;
Best Local Similarity 22.6%; Pred. No. 0.0015;
Matches 73; Conservative 40; Mismatches 100; Indels 110; Gaps 16;

QY 10 QVAGHRASAKGLPLIDSGSLFYKPLQAGDRGEHEVAFYAFSAHAAPVARIIDTFFPRF 69
Db 188 QLAGHTGSGFAAGT-----SGLI---LKRSEPEHYCL-----VRLMADVLRGCVPAF 232
QY 70 HGTRELLPTEAQPGEHPHLVLDLLAGFQAPCVADIKIGAITW----- 112
Db 233 HG-----VVERDGB--SYLQQLDLDGDFGCVLDCKMGVRYLYEELTKARERPKLRKD 285
QY 113 -----PPSP--EPYIAKCLAKDRG-----TTSVLLGFRVSGVRVV--GPEGAVWR 154
Db 286 MYKMLAVDEAPTEEHEAQRVTKRYMQWRREGISSSTTLGFRIGIKKADGSCSTDFK 345
QY 155 TERPEVKAMD-----AGVRVLRVRYSSVADEGMDCALAAAVYGGKGVLSOLRELKAW 209
Db 346 TTRSREQVTRVFEFVQGDDEVLRV-----LNRLQOIRDT 381
QY 210 FEEQTLFHFY---SASILLGVDAAAVAGDGGGVTVKLVDAEHAFAV--GDG-VIDH--- 260
Db 382 LEISDFFRRHEVIGSSLLFVHDCHFRAG-----VWLIDFGKTTPLPDGQILDHRRPW 433
QY 261 -----NFLGGCLSLIKFVSDI 276
Db 434 EEGNREDGYLLGLDNLIGILANL 456

RESULT 3
ARG3_YEAST
ID ARG3_YEAST STANDARD; PRT; 355 AA.
AC P07250;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Arginine metabolism regulation protein III.
GN ARG3 OR ARG82 OR YDR173C OR YD9395.06C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=87257286; PubMed=3239899;
RA Dubois E., Bercy J., Messenguy F.;
RT "Characterization of two genes, ARG1 and ARGRIII required for
RT specific regulation of arginine metabolism in yeast."
RL Mol. Gen. Genet. 207:142-148 (1987).
[2]
SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Murphy L., Harris D.E., Barrell B.G., Rajandream M.A.;
RL Submitted (NCV-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: It is proposed that the ARG1 and ARGRIII proteins are
CC more likely to be involved in the control of ARGRIII activity.
CC -!- MISCELLANEOUS: The ARG1, ARGRII and ARGRIII regulatory proteins
CC are necessary for the repression of six anabolic genes and the
CC induction of two catabolic genes by arginine.
CC -!- MISCELLANEOUS: The expression of this protein is not effected by
CC the presence of arginine.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```



RA Bertsch U., Suesse S., Frerik S., Fanick W.;  
 RT "Cloning of the complete protein coding regions for inositol 1,4,5-  
 RL triphosphate 3-kinase B-isoforms from rat and human."  
 RN Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RP [3]  
 RA SEQUENCE OF 257-946 FROM N.A.  
 RL Donnelly S.;  
 RN Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RP [4]  
 RA SEQUENCE OF 442-946 FROM N.A.  
 RL TISSUE=Brain;  
 RX MEDLINE=91378954; PubMed=1654894;  
 RA Takazawa K., Perret J., Dumont J.E., Erneux C.;  
 RT "Molecular cloning and expression of a new putative inositol 1,4,5-  
 RL triphosphate 3-kinase isoenzyme."  
 RN Biochem. J. 278:883-886 (1991).  
 CC -i- CATALYTIC ACTIVITY: ATP + 1D-myo-inositol 1,4,5-trisphosphate =  
 CC ADP + 1D-myo-inositol 1,3,4,5-tetrakisphosphate.  
 CC -i- ENZYME REGULATION: IP3K is activated by calmodulin. Form B  
 CC is much more sensitive to calcium/calmodulin than form A.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; Y18024; CAB65055.3; -;  
 DR EMBL; AJ242780; CAC40650.1; -;  
 DR EMBL; AL365444; CAD20257.1; -;  
 DR EMBL; X57206; CAA40491.1; ALT\_INIT.  
 DR Genew; HGNC:6179; ITPKB.  
 DR MIM; 147522; -;  
 DR GO; GO:0008440; F:inositol-trisphosphate 3-kinase activity; TAS.  
 DR GO; GO:0007165; P:signal transduction; TAS.  
 DR InterPro; IPR005522; IPK.  
 DR Pfam; PF03770; IPK; 1.  
 KW Transferase; Kinase; Calmodulin-binding.  
 FT DOMAIN 768 776 CALMODULIN-BINDING (BY SIMILARITY).  
 FT POLY-SER.  
 FT CONFLICT 173 173 R -> H (IN REF. 2).  
 FT CONFLICT 210 210 P -> S (IN REF. 2).  
 FT CONFLICT 297 301 GASLT -> ARSEP (IN REF. 2).  
 FT CONFLICT 408 408 A -> S (IN REF. 2 AND 3).  
 FT CONFLICT 442 443 RV -> IP (IN REF. 4).  
 FT CONFLICT 552 552 Q -> P (IN REF. 3).  
 FT SEQUENCE 946 AA; 102391 MW; 36C0C74679B1EA1D CRC64;  
 Query Match 7.2%; Score 109.5; DB 1; Length 946;  
 Best Local Similarity 20.3%; Pred. No. 0.22;  
 Matches 55; Conservative 40; Mismatches 85; Indels 91; Gaps 13;  
 QY 65 PPRPHGTRLLDTEAQGPSEPHVLVDLLAGFQPCVADIKIGAITW-----112  
 Db 711 FVPAYHG-----DVVKDGERVYQ--MDLLADFSCVMDCKMGRTVLEBELTKARKP 763  
 QY 113 -----PPSSP--EPIYAKLAKRG-----TTSVLLGFRVSGVRVVGPEG 150  
 Db 764 SLRKDMYOMIEVDPEAPEEKAQAVTKPRYMQMRETISTATLGFRIEIK--KEDG 821  
 QY 151 AVWRTERPEVKAMDAGVRRVLRVSVSSVADEGMDCALAAVYGGKGVL-----SQREL 206  
 Db 822 TV---NRDFKTKTREQTEAFETK-----GNHLLIAYDRDLKAI 861  
 QY 207 KAMFPEQTLFPHY-----SASILGYDAAVAAGDGGGVTVKLVDFAH---VAEG-----254  
 Db 862 RTTLEVSPPFKCHEVIGSLLFIHDKKEQA-----KWMIDFGKTTPLPEGQTLQHD 913  
 QY 255 ---DGVIDHFLGCLCSLKFVSIVPETP 281  
 Db 914 VPWQEGNREDGYLSGLNLLVILTEMSPQAP 944

RESULT 6  
 ADD MYCLE  
 ID ADD MYCLE STANDARD; PRT; 362 AA.  
 AC Q9CCL9; Q49907;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Adenosine deaminase (EC 3.5.4.4) (Adenosine aminohydrolase).  
 GN ADD OR ML0700 OR L308 C2\_206.  
 OS Mycobacterium leprae  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1769;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Smith D.R., Robison K.;  
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TN;  
 RX MEDLINE=21128732; PubMed=11234002;  
 RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R.M., Deavin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
 RA Holroyd S., Hornaby T., Jagels K., Lacroix C., Maclean J., Moule S.,  
 RA Hurlay L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrell B.G.;  
 RT Massive gene decay in the leprosy bacillus."  
 RL Nature 409:1007-1011 (2001).  
 CC -i- CATALYTIC ACTIVITY: Adenosine + H(2)O = inosine + NH(3).  
 CC -i- SIMILARITY: Belongs to the adenosine and AMP deaminases family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U00022; AAA17330.1; ALT\_INIT.  
 DR EMBL; AL583919; CAC30209.1; -;  
 DR PIR; E86996; E86996.  
 DR HSP; P03958; 1A4M.  
 DR Leproma; MLC700; -;  
 DR HAVAP; MF\_00540; -; 1.  
 DR InterPro; IPR006650; A/AMP deam AS.  
 DR InterPro; IPR001365; A/AMP deaminase.  
 DR InterPro; IPR006330; A deaminase.  
 DR Pfam; PF00962; A deaminase; 1.  
 DR TIGRFAMs; TIGR03430; aden deam; 1.  
 DR PROSITE; PS00485; A\_DEAMINASE; FALSE NEG.  
 KW Hydrolase; Nucleotide metabolism; Complete proteome.  
 FT ACT\_SITE 208 208 POTENTIAL.  
 FT ACT\_SITE 267 267 POTENTIAL.  
 FT ACT\_SITE 300 300 POTENTIAL.  
 FT ACT\_SITE 301 301 POTENTIAL.  
 SQ SEQUENCE 362 AA; 39637 MW; ED0C43B5C92908F5 CRC64;  
 Query Match 6.4%; Score 97.5; DB 1; Length 362;  
 Best Local Similarity 26.8%; Pred. No. 0.73;  
 Matches 56; Conservative 23; Mismatches 91; Indels 39; Gaps 10;  
 QY 83 EPHEHLVLDLLAGFQAP-----CVADIKIGAITWPP--SSPEFYAKLAKDRGT 131  
 Db 69 EFSHTV-----AVMQTFEALHRVAYECVEDLAADSVVYAEVRFAPELHIDGLSFDVL 123  
 QY 132 TSVLLGFRVSGVRVVGPEG-AVWRTERPEVKAMDAGVRRVLRVSVSSVADEGMDCALAA 190



```

|||||
124 ASVLAGEF-ADGERACAEAGNAI---TVRCIVTAMRHAMSRERIAEIAIRFRDKGV---VGF 177
QY 191 AVYGGKGG-----VLSQREILKAMFEQTLFFHYSASILLGYDAAAVAAGDGGGVT 242
Db 178 DIAGAEAGHPTRHDAPEYMRNARF---TIHAGEAFGLPSIHEAIAFCGADRILGHG 233
QY 243 VKLVDFAHVAGSDGVIDHNFLLGGLSLIK 271
Db 234 VRIVDDIDVDPPGGI-----RLGPLASTIR 258

```

## RESULT 7

```

CARA_SULTO
ID CARA_SULTO STANDARD; PRT; 372 AA.
AC Q97008;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Carbamoyl-phosphate synthase small chain (EC 6.3.5.5) (Carbamoyl-
DE phosphate synthetase glutamine chain).
GN CARA OR S11503.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Akai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -!- PATHWAY: Arginine biosynthesis.
CC -!- PATHWAY: Pyrimidine biosynthesis; first step.
CC -!- SUBUNIT: Composed of two chains; the small (or glutamine) chain
CC promotes the hydrolysis of glutamine to ammonia, which is used by
CC the large (or ammonia) chain to synthesize carbamoyl phosphate (By
CC similarity).
CC -!- SIMILARITY: Belongs to the carA family.
CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.

```

```

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----

```

```

EMBL; AF000986; BAB66575.1; -.
DR HAMAP; MF_01209; -.
DR InterPro; IPR006274; CarA synth small.
DR InterPro; IPR001317; CP synthGATase.
DR InterPro; IPR0012474; CP_synthsmall.
DR InterPro; IPR000991; GATase_1.
DR Pfam; PF00988; CPase_sm chain; 1.
DR Pfam; PF00117; GATase; 1.
DR PRINTS; PR00099; CP SGATASE.
DR PRINTS; PR00096; GATASE.
DR TIGRfam; TIGR01368; CPaseIleSmall; 1.
DR PROSITE; PS00442; GATASE TYPE 1; 1.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase;
KW Glutamine amidotransferase; Complete proteome.
FT DOMAIN 1 186

```

```

FT DOMAIN 187 372 GLUTAMINE AMIDOTRANSFERASE.
FT ACT SITE 262 262 GATASE (BY SIMILARITY).
SQ SEQUENCE 372 AA; 41673 MW; 92904F3DCBA3472A CRC64;

Query Match 6.1%; Score 93.5; DB 1; Length 372;
Best Local Similarity 21.7%; Pred. No. 1.6;
Matches 69; Conservative 43; Mismatches 111; Indels 95; Gaps 15;

QY 22 GPLIDGSLFYKPLQAGDGRGEHEVAFYEAFSAHNAVAPARIDTFPRFHGTRLL---PTE 78
Db 18 GTLIEGCGFGAKGIRAG---EVFTSMNGY---PESLTD---PSYRGQILVITHPLV 66
QY 79 AQPGEPPHVLDDLLAGFOAPCVADIKIGAIT-----WPPSSPEPY 120
Db 67 GNGVPEKQRV-EGILTTFES---EQIQVEGLVSEETDFKWNSSKSLHLLWLGSEVPG 122
QY 121 IA---KCLAKDRGTTSVLGFRVGVVVVPGAVNR-----TERPEVKAMD 164
Db 123 LSDVDTRSIYVKVRSRGMVGMVIAAGVEIEDPKKLEKKYDEIDFTQFTSPKAPIVHLGN 182
QY 165 T-----AGVR-----RVLRPY-VSSVADEGMDCALAAAVYGGKGV 200
Db 183 TGTIVVDCGVKHGILYQHORGFTIVRVPCKFNVSKI---MDYYPKGWFGNGFGNP 238
QY 201 SQLRELKAMFEQTLFFHYSASILLGYDAAAVAGDGGGVTVKLVDFAHVAGSDGV 259
Db 239 NLLQVEIENFKELTEYKIPILGICLGHQVATLAWGG-----KVNKKKFGHRAINKPVIDI 293
QY 260 -----HNFLLGGLCS 268
Db 294 SSNKCYITTHNHGYGILS 311

```

## RESULT 8

```

NAGZ_XANCP STANDARD; PRT; 331 AA.
ID NAGZ_XANCP
AC Q8PB42;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Beta-hexosaminidase (EC 3.2.1.52) (N-acetyl-beta-glucosaminidase)
DE (beta-N-acetylhexosaminidase).
GN NAGZ OR XCC1283.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Teal S.M., White F.F.,
RA "Comparison of the genomes of two Xanthomonas pathogens with differing
RA host specificities."
RL Nature 417:459-463(2002).
CC -!- FUNCTION: Cleaves GlcNAc linked beta-1,4 to MurNAc tripeptides (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing N-
CC acetyl-D-hexosamine residues in N-acetyl-beta-D-hexosaminides.

```







```
DE Probable endochitinase precursor (EC 3.2.1.14).
OS Orygia pseudotsugata multicapsid polyhedrosis virus (OpMNPFV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=164623;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97271300; PubMed=9126251;
RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RA Rohrmann G.F.;
RT "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
RT polyhedrosis virus genome.";
RL Virology 229:381-399(1997).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (Potential).
CC -!- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl
CC hydrolases).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U75930; AAC59123.1; -
CC HSSP; P07254; ICTN.
CC InterPro; IPR000886; ER target S.
CC InterPro; IPR001223; Glyco_hydro_18.
CC InterPro; IPR001579; Glyco_hydro_18AS.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR000601; PKD.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC ProDom; PD000471; Glyco_hydro_18; 1.
CC SMART; SM00636; Glyco_18; 1.
CC SMART; SM00089; PKD; 1.
CC PROSITE; PS00014; ER_TARGET; 1.
CC PROSITE; PS01095; CHITINASE_18; 1.
CC Hydrolase; Glycosidase; Chitin degradation; Signal; Glycoprotein;
KW Endoplasmic reticulum.
FT SIGNAL 1 16
FT CHAIN 17 550
FT ACT_SITE 304 304
FT CARBOHYD 146 146
FT CARBOHYD 172 172
FT CARBOHYD 344 344
FT SITE 547 550
FT SEQUENCE 550 AA; 60733 MW; 77947F5CF00E07BD CRC64;
SQ
Query Match 5.8%; Score 87.5; DB 1; Length 550;
Best Local Similarity 21.9%; Pred. No. 8.2;
Matches 68; Conservative 30; Mismatches 105; Indels 107; Gaps 13;
QY 8 EHQVAGHRASAKGLPLDGSGLFYKPLQAGDGRGEHVAFAFAHAHAPARIDTFPP 67
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
75 ENQWKGAAAKRATIDVSKSQFNRRVKLCDD-----DGFVSFPTVRVADT--- 123
QY 68 RPHGTRLLPTE-----AQPGEPPHLY-----LDDLACGQ 98
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
124 --DGHLSPLFVAGENKPKGPHNKTVAAYFVGVGVGRGFFVDKVPPLNLSHLLYGF 181
QY 99 APCVADIKGALTWPSPSPPEPIYAKLAKRGTTSVLLGFRVGVVRVPGCAV--W--- 153
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
182 PICGGDGLNDALKITPGSFALQKSC-----KGRADFKVAIHDPWAAIQPKQGSANNEP 237
QY 154 -----RTERPEVKAMDAG-----VRVLRVYVSVAD----- 181
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
238 YKGNFGQMLAAKLANPHLKIPLSTGGWTLSDPFYFMHDADRRV---FVESVKEFLQWK 294
QY 182 --EGMDCALAAAVYGGKG-----VLSQLRELKAWFEQTL-----PHFVSAS 222
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
-----
DB Db 295 PFDGVD--LDWEFFGKGANPALNGERDADTYLVLLKELRWLDELQLTNKTYELTSA 352
QY 223 ILLGYDAAAV 232
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
353 ISSGYDKIAV 362
RESULT 15
CISZ MYCTU STANDARD; PRT; 373 AA.
AC Q10529;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative citrate synthase 2 (EC 2.3.3.1).
GN CYTA OR RV0899C OR MT0912 OR MTCY31.17C OR MB0913C.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=2206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Usterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating I., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + oxaloacetate = citrate +
CC CoA.
CC -!- PATHWAY: Tricarboxylic acid cycle.
CC -!- MISCELLANEOUS: Citrate synthase is found in nearly all cells
CC capable of oxidative metabolism (By similarity).
CC -!- SIMILARITY: Belongs to the citrate synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
```



1	718.5	47.2	300	10	Q9ARC3	Q9arc3 linum usita
2	714.5	47.0	300	10	Q9FLT2	Q9flt2 arabidopsis
3	704.5	46.3	300	10	Q8LQO5	Q8ldq5 arabidopsis
4	674.5	44.3	286	10	Q8LHG7	Q8lbg7 arabidopsis
5	673.5	44.3	286	10	Q9LYX23	Q9lyx23 arabidopsis
6	664.5	43.7	285	10	Q9ARC1	Q9arc1 linum usita
7	658.5	43.3	286	10	Q9LF72	Q9lf72 arabidopsis
8	288	18.9	309	5	Q5VP86	Q5vpr6 drosophila
9	286	18.8	309	5	Q8MRX9	Q8mrx9 drosophila
10	227	14.9	312	5	Q62519	Q62519 caenorhabdi
11	215.5	14.2	332	5	Q93644	Q93644 caenorhabdi
12	201.5	13.2	416	4	Q8NFU5	Q8nfu5 homo sapien
13	199.5	13.1	371	11	Q8BZA8	Q8bza8 mus musculu
14	199.5	13.1	396	11	Q77T16	Q7tt16 mus musculu
15	194.5	12.8	396	11	Q8BZ11	Q8bz11 mus musculu
16	194	12.8	396	11	Q99NI4	Q99ni4 rattus norv

अ

HGTQSI--EASDGLHPHLILEDLTSTRLHP

	Query Match	47.2%;	Score 718.5;	DB 10;	Length 300;	
	Best Local Similarity	51.6%;	Pred. No. 1.9e-51;			
	Matches 149;	Conservative 44;	Mismatches 83;	Indels 13;	Gaps 7	
Y	7	PEHQVAGHRASAKLGPIIDSGLFYKPLQADGRGEHVAIFYEAFSAHAAPAIRDTRFF	66			
D	5	PEHQVAGHQAINLLGPIVDSDSRRFYKPLQDGERGATEMAYTSPSAHKTIPOHVR-RFF	63			
Y	67	RREHGTRILLPTFAQGCEP-HPHVLVLDLLAGFQPCVADIKIGAITWPPSSPEPYIAKL	125			
	64	PAFHGTQSI--EASDGSLGHPLLIEDITSRLLHPGVMDIKIGRTWTYPFAOAYIEKCL	121			



```
QY 126 AKDRTGTVLLGRVSGVRVVGPE-----GAVWRTERPEVKAMDTAGVRRVRLRVYS--VA 180
Db 122 KKDVESSPFLGRISGLQVYGNCKESSEVVLKPERKLQNLTADEVLVLRKRVSSNPK 181
QY 181 DEGDWDCALAAVYGGKGVLSQRLKAWPEEOTLHFYASILLGV--DAAVAAGDGG 238
Db 182 SDQPDSCFAAVYGGSGNIIQAELLEKANFEDQTIYHFNCSVLMLYEKTKVNGGEE 241
QY 239 --GGVTVKLVDFAHVAGDGVVDHNFGLGCSLIKFSVDIVPETHPTQP 285
Db 242 SLGCAAVKLIIDFAHVTEGNGVDHNFGLGCSLIKFSILTGPDENSP 290

RESULT 2
ID Q9FLT2 PRELIMINARY; PRT; 300 AA.
AC Q9FLT2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Similarity to unknown protein (inositol-(1,4,5) triphosphate
DE 3-kinase) (AT5g61760/mac9_60) (inositol polyphosphate 6-/3-/5-kinase
DE 2b).
GN IP3K OR IPK2B.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98290546; PubMed=9628582;
RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
RT Sequence features of the regions of 1,456,315 bp covered by nineteen
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:41-54(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Xia H.J., Brearley C.A., Mueller-Roeber B.;
RT "Identification of inositol-(1,4,5) triphosphate 3-kinase from
RT Arabidopsis thaliana.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Xia H.J., Brearley C., Mueller-Roeber B.;
RT "Identification of inositol-(1,4,5) triphosphate 3-kinase from
RT Arabidopsis thaliana.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones.";
```

```
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=22302036; PubMed=12226109;
RA Stevenson-Paulik J., Odom A.R., York J.D.;
RT "Molecular and Biochemical Characterization of Two Plant Inositol
RT Polyphosphate 6-/3-/5-Kinases.";
RL J. Biol. Chem. 277:42711-42718(2002).
DR EMBL; AB010069; BAB10076.1; -
DR EMBL; AJ245521; CAC43071.1; -
DR EMBL; AF412073; AAL06526.1; -
DR EMBL; AJ243592; CAC43070.1; -
DR EMBL; AY072621; AAL62012.1; -
DR EMBL; AX147936; AAN63058.1; -
DR GO; GO:000840; F:inositol-trisphosphate 3-kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK, 1.
KW Kinase.
SQ SEQUENCE 300 AA; 33487 MW; 6903A3818CBF27D7 CRC64;
Query Match 47.0%; Score 714.5; DB 10; Length 300;
Best Local Similarity 52.8%; Pred. No. 4e-51;
Matches 150; Conservative 39; Mismatches 82; Indels 13; Gaps 8;
QY 4 LHPPEHQVAGHRASAKLGPLIDGSLGYKPLQAGDGRGHEVAFYAFSAHAAVPARID 63
Db 2 LKVPEHQVAGHTASDGLPLVDQGRFFKPLQDGRSGRGEHAKFYESTSNMKVPDHIH- 60
QY 64 TFEPRFHGTVLLTEAQPGEPPH-PHIVLDDLLAGFOACVADIKIGAITWPPSSPPYIA 122
Db 61 RYFPVYHGTVLV--EASDGSGLKPLHLVDVDSVGYANPSVMDVKIGSRWYPDVSEYFK 118
QY 123 KCLAKDRGTTSVLLGRVSGVRVVG--PEGAVWRTERPEVKAMDTAGVRRVRLRVYS--SV 179
Db 119 KCICKDRQTTVSLGRVSGFKFDHQESSFWAEKKLVGNADGALALRKFVSSNSP 178
QY 180 ADEGM--DCALAAVYGGKGVLSQRLKAWPEEOTLHFYASILLGVDAAAV-AAGG 236
Db 179 ADSNLTPNCAFASVYGGCGNIIQAELLEKDWPEQTQTLHFNSCSILMIYENESILMQGG 238
QY 237 DGG--GVTVKLVDFAHVAGDGVVDHNFGLGCSLIKFSVDIV 277
Db 239 DDAPAPRAQKLVDFAHVLDGNGVDHNFGLGCSFIKFIKDL 282

RESULT 3
Q8LDQ5 PRELIMINARY; PRT; 300 AA.
AC Q8LDQ5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative inositol hexaphosphate kinase.
DE Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085862; AAM63075.1; -
```

```
DR GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR GO:0016301; F:kinase activity; IEA.
DR InterPro: IPR005522; IPK.
DR Pfam: PF03770; IPK; 1.
KW Kinase.
SQ SEQUENCE 300 AA; 33492 MW; DCB7FC6594EF1713 CRC64;
Query Match 46.3%; Score 704.5; DB 10; Length 300;
Best Local Similarity 52.1%; Pred. No. 2.7e-50;
Matches 148; Conservative 40; Mismatches 83; Indels 13; Gaps 8;
QY 4 LHPPEHQVAGHRASAKGLPLDGSGLFYKPLQAGDRGEHEVAFYFAFSAHAAVAPARIRD 63
Db 2 LKYPEHQVAGHIASDGKGLPLVDQGRFFKPLQDGRGEHEAFYFSTNKKVPDHIH- 60
QY 64 TFFPRFHGTRLLPTEAQPGEPH-PLHVLDDLAGFOAPCVADIKIGAITWPPSSPEPIAK 122
Db 61 RYFPVYHGTQLV--EASDGSGLPLHVLDDVSGYANPSPVMDKIGSRITWPDVSEYFK 118
QY 123 KCLAKDRGTTSVLLGFRVSGVRVVG-PEGAVWTERPEVKAMDTAGVRRVLRYYVS--SV 179
Db 119 KCIKDRQXTVSLGFRVSGFKIFDHOESFWAEKKVLVGINADGARLALRKFVSSNP 178
QY 180 ADGEM--DCALAAVYGGKGVLSQRLKAWPEOTLHFYASILLGYDAAAV-AAGG 236
Db 179 ADSNLTPNCAFASEVYGGCGILLAQLELKDWFETQTLYHFNSCSILMIYENESILMKGG 238
QY 237 DGG--GVTVKLVDFHVAEGDGVIDHNFGLGCLSIKIFVSDIV 277
Db 239 DDAAPRAQVQLNFAHVLDDGVNGVIDHNFGLGCLSIKIFKIDIL 282
RESULT 4
Q8LBG7 PRELIMINARY; PRT; 286 AA.
AC Q8LBG7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative inositol hexaphosphate kinase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Arabidopsids II; Brassicales; Brassicaceae; Arabidopsids.
OC eurosids II;
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY087217; AM64773.1; -
DR GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR GO:0016301; F:kinase activity; IEA.
DR InterPro: IPR005522; IPK.
DR Pfam: PF03770; IPK; 1.
KW Kinase.
SQ SEQUENCE 286 AA; 31932 MW; 8351548C2ED8A733 CRC64;
Query Match 44.3%; Score 674.5; DB 10; Length 286;
Best Local Similarity 50.4%; Pred. No. 7.8e-48;
Matches 141; Conservative 42; Mismatches 86; Indels 11; Gaps 7;
QY 4 LHPPEHQVAGHRASAKGLPLDGSGLFYKPLQAGDRGEHEVAFYFAFSAHAAVAPARIRD 63
Db 3 LKYPEHQVAGHIAKDGKPLVDKGRFFKPLQDGRGEHEVAFYFSSNTVEPIH- 61
```

```
QY 64 TFFPRFHGTRLLPTEAQPGEPHVLDDLAGFOAPCVADIKIGAITWPPSSPEPIAK 123
Db 62 RYFPVYHGTQAV--EGSDGA--AMVLENDLAIYSKPSVMDKVGSRITWPDSEYIQK 117
QY 124 CLAKDRGTTSVLLGFRVSGVRVVG-PEGAVWTERPEVKAMDTAGVRRVLRYYVS--SVA 180
Db 118 CLKKDTGTTVSSGFRISGFVYDHKESFWKPERKLRGLDVGARLTILRKVFSSNLS 177
QY 181 DDEM--DCALAAVYGGKGVLSQRLKAWPEOTLHFYASILLGYDAAVAAAGDG 238
Db 178 DTGSKPDSAFASVYGGSHGILLTQLLELKTWFENQTLYHFNSCSILMIYENESILKGNDD 237
QY 239 GG-VTVKLVDFHVAEGDGVIDHNFGLGCLSIKIFVSDIV 277
Db 238 DARPQVKLVDFHVAHVLDDGVNGVIDHNFGLGCLSFNIFIREIL 277
RESULT 5
Q9LY23 PRELIMINARY; PRT; 286 AA.
AC Q9LY23;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T211_80 OR ATSG007370.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
RA Deng J.W., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,
RA Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Shinn P., Tang C.C., Toroumi M., Wallender E.K., Wong C.,
RA Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J.,
RA Theologis A., Davis R.W.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL163912; CAB87926.1; -
DR EMBL: AY136378; AM97044.1; -
DR EMBL: BT000196; AM15515.1; -
DR EIR: T49876; T49876.
DR GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR InterPro: IPR005522; IPK.
DR Pfam: PF03770; IPK; 1.
KW Hypothetical protein.
SQ SEQUENCE 286 AA; 31946 MW; 98578D433C157EFC CRC64;
Query Match 44.3%; Score 673.5; DB 10; Length 286;
Best Local Similarity 50.4%; Pred. No. 9.5e-48;
Matches 141; Conservative 42; Mismatches 86; Indels 11; Gaps 7;
```

```
QY 4 LHPPEHOVAGHRASAKGLPLDGLSGFLFYKPLQAGDRGEHEVAFYFAFSAHAAPARIRD 63
Db 3 LKYPEHOVAGHIAKDGKPGFLVDKGRFFKPLQGDSSRGIEVKFYFSFNTVEPEHIH- 61
QY 64 TFFPRFHGTRLLPTEAQPGEHPHPLVLDLLAGFQAPCVADIKIGAITWPPSPPEPIAK 123
Db 62 RYFPVYHGTQAV--EGSDGA--AMVLENLLAEYTKFSVMDVMKSGRTWYDPDASEEVIQK 117
QY 124 CLAKDRGTTSVLLGFRVSGVRVVG--PEGAVWRTERPEVKAMDTAGVRRVLRVYS--SVA 180
Db 118 CLAKDTGTTVSSGFRISGFEVDHKSFSFKEPKRLRLGLDVGARLTLRKFSVNSLS 177
QY 181 DEGM--DCALAAVYGGKGVLSQRLKAWFEEQTLFHFYSASILLGYDAAVAAGDG 238
Db 178 DTGSKPDSAFASVYGGSHGILTQLELKTWFENQTLHFNSCSILMVYENESILKGNDD 237
QY 239 GG-VTVKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSDIV 277
Db 238 DARPOVKLVDFAHVLDGNGVIDHNFGLGCLSFINFIREIL 277

RESULT 6
Q9ARCI PRELIMINARY; PRT; 285 AA.
AC Q9ARCI;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Linum usitatissimum (flax) (Linseed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Malpighiales; Linaceae; Linum.
OX NCBI_TaxID=4006;
RN [1]
RP SEQUENCE FROM N.A.
RA Dadds P.N., Lawrence G.J., Ellis J.G.;
RT "Identification of the N rust resistance gene of flax and analysis of
RT the role of intragenic sequence exchange in the evolution of the
RT complex N locus.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ310150; CAC35324.1; -.
DR GO; GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK; 1.
KW Hypothetical protein.
SQ SEQUENCE 285 AA; 31542 MW; 8E6A0514100367BF CRC64;

Query Match 43.7%; Score 664.5; DB 10; Length 285;
Best Local Similarity 50.0%; Pred. No. 5.3e-47;
Matches 141; Conservative 45; Mismatches 83; Indels 13; Gaps 7;

QY 4 LHPPEHOVAGHRASAKGLPLDGLSGFLFYKPLQAGDRGEHEVAFYFAFSAHAAPARIRD 63
Db 2 LKYPEHOVAGHIDINGLLGPLVDSDGFRFYKPLQAGDRGATGEMAFYFSFKTIPOXVR- 60
QY 64 TFFPRFHGTRLLPTEAQPGEHPHPLVLDLLAGFQAPCVADIKIGAITWPPSPPEPIA 122
Db 61 RFFQAFQGTQSI--EASDGSGLPHLITLEDITRLHLCVMDIKIGSRTWPEASQAYIE 118
QY 123 KCLAKDRGTTSVLLGFRVSGVRVVG-----PEGAVWRTERPEVKAMDTAGVRRVLRVYS 178
Db 119 KCLAKDVESNPLLGFRISGLQVYNGKSESSEVVLKPERKLLNLNADAEVRLKRFVSS 178
QY 179 -VADEGMDCALAAVYGGKGVLSQRLKAWFEEQTLFHFYSASILLGY--DAAVAAG 235
Db 179 NLKSDQDCSFAAVVSSGNGILLAQLELKAWEEDQTIHFNSCSVLMLEYEKETKWNG 238
QY 236 GDG--GGVTVKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSD 275
Db 239 GEESLGAACKLVDFAHVTEGNGVIDHNFGLGCLSLIKFTISE 280
```

```
RESULT 7
Q9LF72 PRELIMINARY; PRT; 286 AA.
AC Q9LF72;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative inositol hexaphosphate kinase (Inositol polyphosphate
DE 6-/3-/5-kinase 2a).
DE IP6K OR IPK2A.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu J., Mueller-Roeber B., Xu Z.H., Xue H.W.;
RT "no citation.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22302036; PubMed=12226109;
RA Stevenson-Paulik J., Odom A.R., York J.D.;
RT "Molecular and Biochemical Characterization of Two Plant Inositol
RT Polyphosphate 6-/3-/5-Kinases.";
RL J. Biol. Chem. 277:42711-42718(2002).
DR EMBL; AJ404678; CAB96043.1; -.
DR EMBL; AV147935; AAN63057.1; -.
DR GO; GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK; 1.
KW Kinase.
SQ SEQUENCE 286 AA; 31960 MW; 83C47D433C0480FE CRC64;

Query Match 43.3%; Score 658.5; DB 10; Length 286;
Best Local Similarity 49.6%; Pred. No. 1.7e-46;
Matches 139; Conservative 42; Mismatches 88; Indels 11; Gaps 7;

QY 4 LHPPEHOVAGHRASAKGLPLDGLSGFLFYKPLQAGDRGEHEVAFYFAFSAHAAPARIRD 63
Db 3 LKYPEHOVAGHIAKDGKPGFLVDKGRFFKPLQGDSSRGIEVKFYFSFNTVEPEHIH- 61
QY 64 TFFPRFHGTRLLPTEAQPGEHPHPLVLDLLAGFQAPCVADIKIGAITWPPSPPEPIAK 123
Db 62 RYFPVYHGTQAV--EGSDGA--AMVLENLLAEYTKFSVMDVMKSGRTWYDPDASEEVIQK 117
QY 124 CLAKDRGTTSVLLGFRVSGVRVVG--PEGAVWRTERPEVKAMDTAGVRRVLRVYS--SVA 180
Db 118 CLAKDTGTTVSSGFRISGFEVDHKSFSFKEPKRLRLGLYVDGARLTLRKFSVNSLS 177
QY 181 DEGM--DCALAAVYGGKGVLSQRLKAWFEEQTLFHFYSASILLGYDAAVAAGDG 238
Db 178 DTGSKPDSAFASVYGGSHGILTQLELKTWFENQTLHFNSCSILMVYENESILKGNDD 237
QY 239 GG-VTVKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSDIV 277
Db 238 DARPOVKLVDFAHVLDGNGVIDHNFGLGCLSFINFIREIL 277

RESULT 8
Q9VPR6 PRELIMINARY; PRT; 309 AA.
AC Q9VPR6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG13688 protein.
GN CG13688.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
```

[4]  
RN SEQUENCE FROM N.A.  
RP Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[5]  
RN SEQUENCE FROM N.A.  
RP FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AEO03589; AAFA51477.2; --  
DR FlyBase: FBgn0031267; CGI3688.  
DR GO: GO:0008440; Farnesitol-trisphosphate 3-kinase activity; IEA.  
DR InterPro: IPR005522; IPK.  
DR Pfam: PF03770; IPK; 1.  
SQ SEQUENCE 309 AA; 35330 MW; 32G2DDF886094912 CRC64;

Query Match 18.9%; Score 288; DB 5; Length 309;  
Best Local Similarity 30.1%; Pred. No. 9.4e-16;  
Matches 97; Conservative 47; Mismatches 114; Indels 64; Gaps 14

QY	2	SDIHPPPE-----HOVAGH---RASASKGLPIIDG-SGLFYVKPLQAGDGEHEVAIFYEAF	51
Db	4	SDELPEGFROLTKQTAVCHTFFESNAEAVGLLDQSACGVLPKLGKPCGERELRFYESL	63
QY	52	SAHAA-----VPARTDTFFPHGTTRLTLPTAQGPGEHPHLVDLLLAGPQCVCADIK	106
Db	64	AEGASGDNDLLALLRG-HVPRFYGYFLKLVNRR---ERTFLRLDLTRSYAKPCVMGVK	119
QY	107	IGAITW-PPSSP-----EPYIAKCLAKDRGTTSVLGGFRYSVGVRVVGSP	153
Db	120	MGRRTWDPESSNKKEAEKVCKQK-----LGLCLPGFGVYLPKSEHTQETITL	171
QY	154	RTERPEVKAMDTAGVRRVLRVSVSSVADEGM-----CALAAAVYGKGGLVSQRELKAW	209
Db	172	REGKYKSLNVGFKQTMALFNASTSDSKSRAGCELLK-----EVLRLQLQEILAW	225
QY	210	FEEQTLHFYSSAILLYDAAVA-----AGSGGGVTYKLVDPFAHVABGD-	255
Db	226	FQRQLLFHYASSLLICYDSRLADPPKPLNGYHONDDPATVWRVWMIDFAHYVPAEQ	285
QY	256	GVIDHNFLGGLCSLIKFSVDIV 277	
Db	286	GLPDENYMFGLQSLEVVQSIL 307	

RESULT 9  
QBMX9 PRELIMINARY; PRT; 309 AA.

ID	QBMX9	PRELIMINARY;	PRT; 309 AA.
AC	QBMX9;		
DT	01-OCT-2002 (TREMBLrel. 22, Created)		
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)		
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
DE	SD14726P.		
GN	CGI3688.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Nchydroides; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	SEQUENCE FROM N.A.		
RP	Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,		
RA	Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,		
RA	George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,		
RA	Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,		
RA	Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,		
RA	Celniker S.;		
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY119207; AAMS1067.1; --		
DR	FlyBase: FBgn0031267; CGI3688.		
DR	GO: GO:0008440; Farnesitol-trisphosphate 3-kinase activity; IEA.		
DR	InterPro: IPR005522; IPK.		
DR	Pfam: PF03770; IPK; 1.		
SQ	SEQUENCE 309 AA; 35312 MW; 7EA209AC834C5EAF CRC64;		

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Phyloidea; Drosophilidae; Drosophila.  
[1]  
RN NCBI\_TaxID=7227;  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RX MEDLINE=20196006; PubMed=107311132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,  
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Beriman B.P., Bhargava D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,  
RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foeller C., Gabrielson A.E., Gang N.S., Galbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan K.A.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M.C., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
[2]  
RN SEQUENCE FROM N.A.  
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Anantides P.G., Brandon R.C., Rogers Y.,  
RA Hanson J., An H., Baldwin D., Banon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwan K.A., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
RT "Sequencing of Drosophila melanogaster genome.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[3]  
RN SEQUENCE FROM N.A.  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradscky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
RA Tupay J.L., Bergman B., Carlson J.W., Celniker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Seattle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.S., Rubin G.M., Mungall C.J., Lewis S.E.;  
RT "Annotation of Drosophila melanogaster genome.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.



```
Q8NFUS PRELIMINARY; PRT; 416 AA.
AC Q8NFUS;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Inositol polyphosphate multikinase.
GN IPMK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22176314; PubMed=12027805;
RA Nalauskowski M.M., Deschermeier C., Panick W., Mayr G.W.;
RT "The human homologue of yeast ArgR111 protein is an inositol phosphate
RT multikinase with predominantly nuclear localization.";
RL Biochem. J. 366:549-556(2002).
DR EMBL; AF432853; AAM97838.1; -
DR GO; GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK; 1.
KW Kinase.
SQ SEQUENCE 416 AA; 47222 MW; 622D678696A892F9 CRC64;
Query Match 13.2%; Score 201.5; DB 4; Length 416;
Best Local Similarity 28.4%; Pred. No. 2e-08;
Matches 71; Conservative 34; Mismatches 110; Indels 35; Gaps 9;
QY 6 PPEHQVAGHRASAKGLPLIDSGLFYKPLQAGDGRGEHEVAFYE---AFSAHAAVPAIR 62
DB 48 PLSHQVAGHMYGKQKVGLIQLHPDGVTLKQLQPPRGPRELEFVYMYAACDCGVLLELR 107
QY 63 DTFPRFHGTRLLPTEAQGEHPHVLVDLLAGFOAPCVADIKIGAITWPPSSPEPYIA 122
DB 108 -KYLKYGIWSPPT--APND--LVLKLEDTVTHKPKCINDVKIGKSYDPPASSEKIQ 162
QY 123 KCLAKDRGTTSVLLGPRVSGVRVVGEGAVWTE-----RPEVKAMDTAGVRVLRYS 177
DB 163 QQVSKYPLMEEI--GFLVLGMRYVHLSDSYETQNHQYGRSLTKETIKDGVSRFFHNGYC 220
QY 178 SVADGMDCALAAVYGGKGVLSQLRELKAWFEETLHFYSASILLGYDAAVAAGD 237
DB 221 LRKD-----AAVASI-----QKTEKLQWFENKQLNFYASSLLFYFE-----GS 260
QY 238 GSGVTIVKLVD 247
DB 261 SQFTTKLND 270
RESULT 13
Q8BZA8 PRELIMINARY; PRT; 371 AA.
AC Q8BZA8;
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Inositol polyphosphate MULTIKINASE homolog (Fragment).
GN IMPK OR 2410017C19IK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
```

```
RL Nature 420:563-573(2002).
DR EMBL; AK036084; BAC29300.1; -.
DR PIR; PT0645; PT0645.
DR MGI; MGI:1916968; Impk.
DR GO; GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK; 1.
FT NON_TER 1
SQ SEQUENCE 371 AA; 41888 MW; 1C9529B05DA34DF5 CRC64;
Query Match 13.1%; Score 199.5; DB 11; Length 371;
Best Local Similarity 26.3%; Pred. No. 2.6e-08;
Matches 70; Conservative 47; Mismatches 122; Indels 27; Gaps 10;
QY 6 PPEHQVAGHRASAKGLPLIDSGLFYKPLQAGDGRGEHEVAFYEAFSAHAAVPAIRD-- 63
DB 6 PLSHQVAGHMYGKQKVGLIQLHPDGVTLKQLQPPRGPRELEFVYMYAACDADAVLELR 65
QY 64 TFFPRFHGTRLLPTEAQGEHPHVLVDLLAGFOAPCVADIKIGAITWPPSSPEPYIAK 123
DB 66 KHLPKYGIWSPPT--APND--VYLKLEDTVTHKPKCINDVKIGKSYDPPASSEKIQQ 121
QY 124 CLAKDRGTTSVLLGPRVSGVRVVGEGAVWTEPEVKAMDTAGVRVLRVYSSVADEG 183
DB 122 QVSKYPLMEEI--GFLVLGMRYVHLSDSYETQNHQYGRSLT---KETLKEGYSKFFHNG 176
QY 184 MDC---ALAAVYGGKGVLSQLRELKAWFEETLHFYSASILLGYDAAVAAGDGG 239
DB 177 F-CLRKDAVAASI-----QKTEKLQWFENKQLNFYASSLLFYVEGSGQAPATKAN 227
QY 240 GVTIV--KLVDFAHVAEGDGV-IDHNF 262
DB 228 DRTLGRFLSKGLPLTDADGLECNMF 253
RESULT 14
Q7TT16 PRELIMINARY; PRT; 396 AA.
AC Q7TT16;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC052463; AAH52463.1; -.
KW Hypothetical protein.
SQ SEQUENCE 396 AA; 44453 MW; 5A91EED2B601E977 CRC64;

Query Match      13.1%; Score 199.5; DB 11; Length 396;
Best Local Similarity 26.3%; Pred. No. 2.8e-08;
Matches 70; Conservative 47; Mismatches 122; Indels 27; Gaps 10;

QY 6 PPEHQVAGHRASAKGLPLDGSGLFYKPLQAGRGHEVAFAFSAHAAPARIRD-- 63
D 31 PLSHQVAGHYMKDKVGLQHPDGVTLKQLQPPRGPRGPRLEFFYTMVYAADCADAVLLELR 90
QY 64 TFFPRFHGTRLPTAOPGEPHPLVLDLLAGFOAPCVADIKIGAITWPPSSPEPYIAK 123
D 91 KHLPKYGVWSPPPT--AFND--VYLKLEDTVHKENKPCIMDVKIGRKSYDPFASSEKIQQ 146
QY 124 CLAKDRGTTSVLLGFRVSGVRVVGPEGAVWRTPEVKAMDTAGVRRVLRVYSSVADEG 183
D 147 QVSKYPLMEEI--GFLVLGMRVYHLHSDSYETQNHYGRGLT---KETLKEGVSKFFHNG 201
QY 184 MDC-----ALAAVYGGKGVLSQLRELKAWFEEQTLFHFYSASILLGYDAAVAAGDGG 239
D 202 F-CIRKGAIAASI-----QKVEKILQWFENQKQLNFYASSLLFVYEGSSQPATTKAN 252
QY 240 GVTV--KLVDFAHVAEGDGV-IDHNF 262
D 253 DRTLGRFLSKGPLTDADGLECENN 278

Search completed: March 24, 2004, 20:59:08
Job time : 48 secs

RESULT 15
Q8BZ11
ID Q8BZ11 PRELIMINARY; PRT; 396 AA.
AC Q8BZ11;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Inositol polyphosphate MULTIKINASE homolog.
GN IMPK OR 241017C19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Vagina;
RX MEDLINE=22354683; PubMed=1246851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RT Nature 420:563-573(2002).
DR EMBL; AK036978; BAC29655.1; -.
DR FIR; PT0645; PT0645.
DR MGD; MGI:1916968; Impk.
DR GO; GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK; 1.
SQ SEQUENCE 396 AA; 44410 MW; 26F9699A4C680B0A CRC64;

Query Match      12.8%; Score 194.5; DB 11; Length 396;
Best Local Similarity 26.3%; Pred. No. 7.3e-08;
Matches 70; Conservative 47; Mismatches 122; Indels 27; Gaps 10;

QY 6 PPEHQVAGHRASAKGLPLDGSGLFYKPLQAGRGHEVAFAFSAHAAPARIRD-- 63
D 31 PLSHQVAGHYMKDKVGLQHPDGVTLKQLQPPRGPRGPRLEFFYTMVYAADCADAVLLELR 90
QY 64 TFFPRFHGTRLPTAOPGEPHPLVLDLLAGFOAPCVADIKIGAITWPPSSPEPYIAK 123
D 91 KHLPKYGVWSPPPT--AFND--VYLKLEDTVHKENKPCIMDVKIGRKSYDPFASSEKIQQ 146
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 27, 2004, 05:59:59 ; Search time 3160 Seconds  
(without alignments)  
2731.067 Million cell updates/sec

Title: US-10-042-894A-8  
Perfect score: 1521  
Sequence: 1 MSDLLPPEHQVAGHRASAK.....IKFVSDIVPETHPTQLGPS 289

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO spoal/US10042894/runat\_24032004\_145047\_16959/app\_query.fasta\_1.455  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HSPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10042894 @CGN 1.1 3437 @runat\_24032004\_145047\_16959 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:

1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hcc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hcc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rtd.\*  
26: em\_gss\_pbg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1521	100.0	960	28	CC344833	CGIAQ20TV
2	1406.5	92.5	923	29	CG240397	CGYCI27TV
3	1256	82.6	818	29	CG453578	CG9AD40TV
4	1248	82.1	873	28	CC344824	CGIAQ20TH
5	1183	77.8	867	29	CG724960	CGLAB23TV
6	1149	75.5	719	29	CG284572	CGWIF61TH
7	1123	73.8	801	13	CA134480	SCJFRT106
8	1034	68.0	656	14	CA245555	SCBGF508
9	993.5	65.3	722	13	CA130685	SCCCT100
10	992	65.2	591	29	CG776236	1123005E0
11	987	64.9	671	14	CA202412	SCRLF100
12	981	64.5	3374	11	AY109355	Zea mays
13	976	64.2	722	29	CC724950	OGLAB23TH
14	932	61.3	744	13	CA130686	SCCCT100
15	926	60.9	1191	14	CK211413	FGAS02325
16	921.5	60.6	593	28	BZ774817	i151a11.b
17	920	60.5	960	29	CG073495	FUJCG54TD
18	903	59.4	955	28	CC384766	FUHO53TD
19	869.5	57.2	794	10	BF617713	HVSMBC001
20	811.5	53.4	907	29	CG240387	OGYCI27TH
21	807	53.1	575	29	CG734025	1119161G0
22	770	50.6	619	14	CA204213	SCAGFL109
23	758	49.8	1099	14	CK215908	FGAS02788
24	727	47.8	1123	14	CK214494	FGAS02642
25	724.5	47.6	588	13	B0999708	HI15H04r
26	724	47.6	910	14	CK253394	EST773031
27	714	46.9	551	28	BH408067	1007056A0
28	644	42.3	557	10	BF277833	GA_Eb003
29	642	42.2	831	29	CG953151	MBEFW65TF
30	640	42.1	620	13	BQ620191	TaLr1176F
31	613	40.3	453	12	B1956405	HVSMEN000
32	606	39.8	433	12	BM099879	EBES01_SQ
33	584	38.4	448	28	CC028437	3591_1_10
34	581.5	38.2	478	12	BG417083	HVSNK001
35	580	38.1	876	29	CG073494	FUJCG54TB
36	567	37.3	409	12	BM099878	EBES01_SQ
37	562	36.9	927	29	CG284580	OGWIF61TV
38	559	36.8	431	29	CG776505	1123005E0
39	529	34.8	317	29	CG194403	FUJGW38TB
40	527	34.6	767	29	CG971224	MBEF51TF
41	527	34.6	823	29	CG931827	MBSCW31TF
42	522	34.3	688	13	BQ405077	GA_EB007
43	517	34.0	739	14	CA917416	EST641563
44	508.5	33.4	629	13	B0993995	HM0521r
45	496	32.6	691	28	BH585814	BOGHV81TF

# ALIGNMENTS

RESULT 1  
CC344833  
LOCUS  
DEFINITION CG344833 CGIAQ20TV ZM 0.7.1.5 KB Zea mays genomic clone ZMMBMA0359C15, linear GSS 16-MAY-2003  
genomic survey sequence.  
ACCESSION CC344833  
VERSION CC344833.1 GI:30814239  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Euphorbiaceae; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.



```
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 34 ATGCCCGACCTCCACCCGCGGAGCACCAAGTCGCGGTTCACCGGCTCCGCCAGCAG 93
QY 21 LeuGlyProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 94 CTGGGCGCGCTCATCGACGCGCTCGGCGCTCTTCTACAGCGGCTCCAGGCGCGACCGT 153
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProLys 60
Db 154 GGGGAGCAGAGTCGCTCTTATAGGCGGTCTCCGCCACGCGCGCTCCGCCCGC 213
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 214 ATCCGAGACACCTCTCTCCCGCGTTCAGGCGCGCTCTCCGCCAGGCGCGAG 273
QY 81 ProGlyGluProHisProHisLeuValLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
Db 274 CCGGGGAGCGCGCATCCGACCTCGTCTCGACGACCTCTCGCGGGTTTGAGGCGCC 333
QY 101 CysValAlaAspTLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
Db 334 TCGCTCGCAGACATCAAGATCGCGCGCATCAGTGGCCACCGAGTTCGCGGAGCGCTAC 393
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 394 ATCCCAAGTACTCGCCAGGACCGCGGAGCACCGAGCGTCTCGGTCTCGGTTCGCGTC 453
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
Db 454 TTGC--GTCCGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 511
QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-Val 180
Db 512 AAGGCTATGGACACCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 571
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyVal 200
Db 572 CGACGAGGGATGGAGCTGCGCGCTCGCGCGCGCGGTGTACGAGGAGAAAGTGGAGTCT 631
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 632 GTACAGCTGCGCGAGCTCAAGGCTGTTGGAGGAGCAGACTGTCTTCTTCTTCTTCT 691
QY 220 rAlaSerLeuLeuGlyTyrAspAlaAlaAlaValTyrGlyGlyLysGlyGlyVal 240
Db 692 GCGCTCGATCTCTCGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 751
QY 240 YValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 752 GGTACAGTGAAGCTGGTGGACTTTGCCCATGTGCCGAGGCTGATGGGTTGATTGACCA 811
QY 260 sAnPheLeuGlyGlyLeuCysSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 280
Db 812 CAACTTCTGGCGAGCTCTGCTAGTGTATCAAGTTCGTTCTGACATTTGTTCCAGAGAC 871
QY 280 rProHisThrGlnProLeuGlyProSer 289
Db 872 TCTTAGCGGAGCTTTGGGCTCTTCT 899

RESULT 3
CG453578
LOCUS
DEFINITION
  OG9AD40TV_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBMA0849H08,
  genomic survey sequence.
ACCESSION
  CG453578
VERSION
  CG453578.1
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 818)  
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other GSSs: OG9AD40TH  
Contact: Cathy Whitelaw

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TF  
Class: sheared ends.

FEATURES  
source

1..818  
/location/Qualifiers  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0849H08"  
/clone\_lib="ZM 0.7 1.5 KB"  
/note="Vector: pBCKS-; site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:

Pred. No.: 5,92e-112 Length: 818  
Score: 1256.00 Matches: 239  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.58% Indels: 0  
DB: 29 Gaps: 0

US-10-042-894a-8 (1-289) x CG453578 (1-818)

```
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 96 ATGTCGCGACCTCCACCCGCGGAGCACCAAGTCGCGCGCTCCGCCAGCAG 155
QY 21 LeuGlyProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 156 CTGGGCGCGCTCATCGACGCGCTCTTCTACAGCGCTCCAGGCGCGACCGT 215
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProLys 60
Db 216 GGGGAGCAGAGTCTCTTATGAGCGGTCTCCGCCAGCGCGCTCCGCCCGC 275
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 276 ATCCGAGACACCTCTTCTCCCGGTTCAGGCGCGCTCTCCGCCAGGCGCGAG 335
QY 81 ProGlyGluProHisProHisLeuValLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
Db 336 CCGGGGAGCGCGCATCTCTACCTCGTCTCGACGACCTCTCGCGGGTTTCAGGCGCC 395
QY 101 CysValAlaAspTLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
Db 396 TCGCTCGCAGACATCAAGATCGGCGCATCAGTGGCGCGCGCGCGCGCGCGCG 455
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 456 ATCCCAAGTCTCCGCCAGGACCGCGGAGCGCGGTCTGCTCGGATTCGCGCTC 515
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
Db 516 TCGGCGCTCCGAGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 575
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgValLeuArgTyrValSerSerVal 180
Db 576 AAGGCTATGGACACCGCGCGCGTCCGCGCGGTCTCGCGCGCTAGTGTCTATCGTTC 635
```

```

QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyGlyValLeu 200
Db 636 GAGCAGGGATGACTGTGGCTCGCGCGCGGTGTACGAGGAAAAGGTGGAGTCTTG 695
QY 201 SerGlnLeuArgGluLeuLysAlaTTPheGluGlnThrLeuPheHisPheTyrSer 220
Db 696 TCACAGCTGGCAGACTCAAGGGGTGGTTCAGAGCAGACTCTGTCCACTTCTACTCG 755
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 239
Db 756 GCGTCGATTCTTCTGGCTATGATGCTGTGCAAGTCCGACAGCGGAGATGGGGGG 812

RESULT 4
CC344824/c 873 bp DNA linear GSS 16-MAY-2003
LOCUS OGIAQ020TH ZM 0.7.1.5 KB Zea mays genomic clone ZMBMa0359C15,
DEFINITION genomic survey sequence.
ACCESSION CC344824
VERSION CC344824
KEYWORDS CC344824.1 GI:30814230
SOURCE GSS.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1. (bases 1 to 873)
White, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1. .873
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBMa0359C15"
/clone_lib="ZM 0.7.1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.: 3,94e-111 Length: 873
Score: 1248.00 Matches: 238
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 82.05% Indels: 0
DB: 28 Gaps: 0

US-10-042-894A-8 (1-289) x CC344824 (1-873)

QY 52 SerAlaHisAlaValProAlaArgIleArgAspThrPheProArgPheHisGly 71
Db 872 TCCGCCACACCGCGCTCCGCGCGCATCCGACACCTTCTTCCCGCGTTCACCGC 813
QY 72 ThrArgLeuLeuProThrGluAlaGlnProGlyGluProHisProHisLeuValLeuAsp 91
Db 812 ACGGACTCTTCCCCAGGGCGCAGCCCGGGAGCCGATCTTCACTTCGTCCTCGAC 753
QY 92 AspLeuLeuAlaGlyPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThr 111
Db 752 GACTCTCTCGGGGGTTTCAGCGCCCTCGCTCGCAGACATCAAGATCGCGCCATCAG 693

```

```

QY 112 TrpProProSerSerProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThr 131
Db 692 TGGCCACCCAGATTGCGCGAGGCCCTACATCGCAAGTGCCTGCCAAGGACCGCGGACC 633
QY 132 ThrSerValLeuLeuGlyPheArgValSerGlyValArgValValGlyProGluGlyAla 151
Db 632 ACAGAGGTTCTGCTCGGATTCGCGTCTCCGGGTCGAGTGTGCGGCCCGGAGGGCGCC 573
QY 152 ValTrpArgThrGluArgProGluValLysAlaMetAspThrAlaGlyValArgArgVal 171
Db 572 GTGTGGCGACGAGCGCCCGAGGTGAAGGCATGACACCGCGCGCTCCGCGCGCTG 513
QY 172 LeuArgArgTyrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAla 191
Db 512 CTCGCGCGCTAGTGTGTCATCCGTCGCGAGGGGATGACTGTGCTCGCGCGCGC 453
QY 192 ValTyrGlyGlyLysGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGlu 211
Db 452 GTGTACGAGGAAAAGGTGGAGTCTTGTCCAGCTCCGCGAGCTCAAGGGGTGGTTCGAG 393
QY 212 GluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAla 231
Db 392 GAGCAGACTCTGTCCACTTCTACTCGCGGTGCGATTCTTCTGGGCTATGATGCTGCTGCA 333
QY 232 ValAlaAlaGlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisVal 251
Db 332 GTCGACAGCGGAGGATGGGGTGGGGTGGAGCTGAGCTGAGCTGAGCTGATGCTGCTG 273
QY 252 AlaGluGlyAspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLys 271
Db 272 GCGAGGGTGTGGGGTGAATGACCAACTTCTCTGGGGGGGCTCTGCTGCTGATCAAG 213
QY 272 PheValSerAspIleValProGluThrProHisThrGlnProLeuGlyProSer 289
Db 212 TTCTGTTTCTGACATTGTTCCGGAGAGCTCCTCATACGACGCTTTGGTCTCTTCT 159

RESULT 5
CC724960 867 bp DNA linear GSS 23-JUN-2003
LOCUS OGLAB23TV ZM 0.7.1.5 KB Zea mays genomic clone ZMBMa0307C22,
DEFINITION genomic survey sequence.
ACCESSION CC724960
VERSION CC724960.1 GI:32143893
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1. (bases 1 to 867)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGLAB23TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
source
1. .867
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBMa0307C22"
/clone_lib="ZM 0.7.1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb"

```

## methylation filtered genomic DNA library"

## ORIGIN

## Alignment Scores:

Pred. No.: 8,54e-105 Length: 867  
Score: 1183.00 Matches: 236  
Percent Similarity: 90.15% Conservative: 2  
Best Local Similarity: 89.39% Mismatches: 7  
Query Match: 77.78% Indels: 20  
DB: 29 Gaps: 1

US-10-042-894A-8 (1-289) x CC724960 (1-867)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
DB 131 ATGCCGACCTCCACCGCGGAGCACCAAGTCGCGGTCAACGCGCTCCGCGCAAG 190  
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
DB 191 CTGGCCCACTCATCGAGACTCTGGCTCTTCTACAGCGCTCCAGCGCGGACCGT 250  
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
DB 251 GGGGAGCAGCGAGTCTCTATAGGCGTCTCCGCCACCGCGCGTCCGCGCGCG 310  
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
DB 311 ATCCGAGACACTTCTTCCCGCGGTCCACGGACGAGCTCTCCACCGAGGCGAG 370  
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100  
DB 371 CCGGGGAGCGCATCGCACCTCTCTCGACGACTCTCTCGGGGTTTGAGGCGCC 430  
QY 101 CysValAlaAspIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
DB 431 TCGCTCGCAGACATCAAGATCGGTGCCATCACGTG----- 465  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
DB 466 -----ACCACGAGCGTTCTGCTCGGATTCCGCGTC 495  
QY 141 SerGlyValArgValValGlyProGluGlyValAlaValTyrArgThrGluArgProGluVal 160  
DB 496 TCCGGCTCCGACTCTCGGCCCGGAGGCGCGTGTGGCGAGCGAGCGCGCGAGGTG 555  
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180  
DB 556 AAGGCTATGGACATTCTCGCGCTCCGCGCTCTCCGCGCTTACGTGTATCCGCTGC 615  
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyGlyGlyValle 200  
DB 616 CGACGAGGGGATGGACTGCGCTCGCGCGCGGTGTACGAGGAAAGGTGGATCTT 675  
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220  
DB 676 GTCACAGCTGCGGAGCTCAAGCGGTGTTCGAGGGGCGAGCTCTCTTCCACTTCTACTC 735  
QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240  
DB 736 GCGCTCGATTCTTCTGGCTATGATCTGTGTCAGTCGACGAGCGAGGTGGGGTGG 795  
QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260  
DB 796 GGTAAAGTGAAGCTGGTGAATTTGCCATGTGGCGAGGTGATGGGGTGTATTACCA 855  
QY 260 sAenPheLeu 263  
DB 856 CAACITCTCTG 865

## RESULT 6

CG284572 719 bp DNA linear GSS 25-AUG-2003  
LOCUS  
DEFINITION OGWIF61TH ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0584K02,

## ACCESSION

CG284572  
VERSION CG284572.1 GI:34198786  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

## REFERENCE

AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.  
TITLE Consortium for Maize Genomics  
JOURNAL Unpublished (2002)  
COMMENT Other GSSs: OGWIF61TV

## FEATURES

source Location/Qualifiers  
1..719  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0584K02"  
/clone\_lib="ZM\_0.7\_1.5\_KB"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"  
TIGR  
Contact: Cathy Whitelaw  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

## FEATURES

source Location/Qualifiers  
1..719  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0584K02"  
/clone\_lib="ZM\_0.7\_1.5\_KB"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.33e-101 Length: 719  
Score: 1149.00 Matches: 230  
Percent Similarity: 89.92% Conservative: 2  
Best Local Similarity: 89.15% Mismatches: 7  
Query Match: 75.54% Indels: 20  
DB: 29 Gaps: 1

US-10-042-894A-8 (1-289) x CG284572 (1-719)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
DB 3 ATGCCGACCTCCACCGCGGAGCACCAAGTCGCGGTCAACGCGCTCCGCGCAAG 62  
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
DB 63 CTGGGCGGACCTCATCGACGACTCTGGCTCTTCTACAGCGCTCCAGCGCGGACCGT 122  
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
DB 123 GGGGAGCAGAGGTGCGCTTCTATAGGCGTCTCCGCCACCGCGCGCTCCGCGCGCG 182  
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
DB 183 ATCCGAGACACTTCTTCCCGCGGTTCACGCGACGCGACTCTCTCCACCGAGGCGAG 242  
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100  
DB 243 CCGGGGAGCGCATCCGACCTCTCTCTCGACGACTCTCTCGGGGTTTGAGGCGGCC 302  
QY 101 CysValAlaAspIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
DB 303 TCGCTCGCAGACATCAAGATCGGTGCCATCACGTG----- 337  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
DB 338 -----ACCACGAGGTTCTGCTCGGATTCGCGTC 367

```

QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrrArgThrGluArgProGluVal 160
Db 368 TCCGGCGTCCAGTTCGTGGCGCCCGAGGGCCCGTGTGGCGACGAGCGCCCGAGGTG 427
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTrrValSerSer-ValAl 180
Db 428 AAGGCTATGGACATGTGGCGTCCGCGCGTGTCCGGCGCTACGTGTATCCGCTTGC 487
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTrrGlyGlyGlyGlyValle 200
Db 488 CGACGAGGGATGACTGCGCGTCCGCGCGCGTGTACGAGCAAAAGGTGGAGTCTT 547
QY 200 uSerGlnLeuArgGlnLeuLysAlaTrrPheGluGlnThrLeuPheHisPheTrrSe 220
Db 548 GTACAGCTGGCGAGCTCAAGGCGTGTTCGAGGGCGAGACTGTGTCCACTTCTACTC 607
QY 220 rAlaSerIleLeuLeuGlyTrrAspAlaAlaAlaValAlaAlaGlyGlyGlyGly 240
Db 608 GGCGTGGATTCCTTGGGCTATGATGTGTGTGCTGCTGAGTGCAGAGCGAGGTGGGGTGG 667
QY 240 YValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
Db 668 GSTAACAGTGAAGCTGTGGACTTGGCCATGTGGCGGAGGTGATGGGTG 719

RESULT 7
LOCUS CAL134480 801 bp mRNA linear EST 24-SEP-2003
DEFINITION SCJFRT1061H11.9 RT1 Saccharum officinarum cDNA clone SCJFRT1061H11
5', mRNA sequence.
ACCESSION CAL134480
VERSION CAL134480.1 GI:35021536
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.
1 (bases 1 to 801)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 061 row: H column: 11
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .801
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFRT1061H11"
/lab_host="DH10B"
/clone_lib="RT1"
/note="Organ: Root tips (0.3cm-long) from adult plants;
Vector: pSport1; Site_1: SalI; Site_2: NotI; An
unidirectional cDNA library generated from [Root tips
(0.3cm-long) from adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

```

## ORIGIN

Alignment Scores:  
 Pred. No.: 5,39e-99 Length: 801  
 Score: 1123.00 Matches: 220  
 Percent Similarity: 89.58% Conservative: 12  
 Best Local Similarity: 84.94% Mismatches: 24  
 Query Match: 73.83% Indels: 3  
 Gaps: 13

US-10-042-894A-8 (1-289) x CA134480 (1-801)

```

QY 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 31 ATGTCCGACCTCCGCCCGCGGAGCACCAAGTCGCGGCGCACCGCGCTCCGCCAACAG 90
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTrrLysProLeuGlnAlaGlyAspArg 40
Db 91 CTGGGTCCGCTCATCGACGGCTCTGGCTCTTCTACAGCGCTTCCACCCACCGCGCTCCGCGCCGC 150
QY 41 GlyGluHisGluValAlaPheTrrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 151 GGGGAGCAGCAGCTCGCTTCTACGAGGCGTTCACACCCACCGCGCTCCGCGCCGC 210
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 211 ATCCGGGACACCTCTTCCCGCGGTTCACGGCAGCGACTCTCTCCACCGAGGCGCG 270
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 271 CCAGGGGAGCGCATCCGACCTCTCTCGACGACTCTCTCGCGGGCTGGAGGCGCC 330
QY 101 CysValAlaAspIleLysIleGlyAlaIleThr-TrrProProSerSer-ProGluProT 120
Db 331 TGGCTCGCGACATCAAGATCGGCGCATCACAGTGGCGCGCGAGCTCGCGCAGAGCCCT 390
QY 120 YrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgV 140
Db 391 ACCTGTCAAGTGCCTGGTCAAGGACCGCGGACCGAGCATTCCTGCTCGGATTCGCG 450
QY 140 alSerGlyValArgValValGlyProGluGlyAlaValTrrArgThrGluArgProGluV 160
Db 451 TCTCCGGGCTCCGGGTCTCGTCCGAGGGCGCGCTGTGGCGAGCTGACGCGCGGAGG 510
QY 160 alLysAlaMetAspThrAlaGlyValArgValLeuArgTrrValSerSerValAla 180
Db 511 TGAAGCGCTGGACACCGCGCGTCCGCGCTGTCTCCGCGCTAGCTTTCATCCGTTG 570
QY 180 laAspGluGlyMetAspCysAlaLeuAlaAlaValTrrGlyGlyLysGlyVal 200
Db 571 CCAGCAGGGGATGGACTGCGACTCGCGCGCGGTGTACGGCGGNAAGGGAGTTT 630
QY 200 euSerGlnLeuArgGlnLeuLysAlaTrrPheGluGlnThrLeuPheHisPheTrrS 220
Db 631 TGTCAACAATTGGCGAGCTTAAGCGTGTGTTCGAGGAACAGACTCTGTGTCCACTTACT 690
QY 220 erAlaSerIleLeuLeuGlyTrrAspAlaAlaValAlaAlaGlyGly-AspGlyGly 239
Db 691 CGGGGGCGATTCTTCTGGGCTATGAAGCTAGTCCGCTTGCACCAACCGGAGGTAAGGT 750
QY 240 GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGly 256
Db 751 GGGGTGAAGGTAAACTGGGGGCCCTTTCGCATGTGGCCCAAGGGAAGGG 801

```

## RESULT 8

CA245555 656 bp mRNA linear EST 25-SEP-2003  
 SCBGFL5081A03.g Saccharum officinarum FL5 Saccharum officinarum  
 LOCUS cDNA clone SCBGFL5081A03 5', mRNA sequence.  
 DEFINITION  
 ACCESSION CA245555  
 VERSION CA245555.1 GI:3524301  
 KEYWORDS EST.  
 SOURCE Saccharum officinarum

ORGANISM Saccharum officinarum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Saccharum.  
1 (bases 1 to 656)  
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bcccenter.fcav.unesp.br  
Plate: 081 row: A column: 03  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers  
1..656  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCBGL5081A03"  
/lab\_host="DH10B"  
/clone\_lib="Saccharum officinarum FL5"  
/notes="Organ: Developed inflorescence (20cm-long) without  
rachis; Vector: pSport1; Site 1: SalI; Site 2: NotI; An  
unidirectional cDNA library generated from [Developed  
inflorescence (20cm-long) without rachis]. cDNA was  
prepared from polyA+ mRNA using SuperScript Plasmid  
System Kit (Invitrogen). The double-strand cDNAs were  
fractionated in a sepharose CL-2B 40cm-columns and  
fragments sizing between 0.8 and 1.5 Kb were  
directionally cloned into the vector. Details of each  
source of RNA and library construction can be obtained at  
http://sucest.lad.ic.unicamp.br/public"

FEATURES  
source  
1..656  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCBGL5081A03"  
/lab\_host="DH10B"  
/clone\_lib="Saccharum officinarum FL5"  
/notes="Organ: Developed inflorescence (20cm-long) without  
rachis; Vector: pSport1; Site 1: SalI; Site 2: NotI; An  
unidirectional cDNA library generated from [Developed  
inflorescence (20cm-long) without rachis]. cDNA was  
prepared from polyA+ mRNA using SuperScript Plasmid  
System Kit (Invitrogen). The double-strand cDNAs were  
fractionated in a sepharose CL-2B 40cm-columns and  
fragments sizing between 0.8 and 1.5 Kb were  
directionally cloned into the vector. Details of each  
source of RNA and library construction can be obtained at  
http://sucest.lad.ic.unicamp.br/public"

ORIGIN  
Alignment Scores:  
Pred. No.: 1,91e-90 Length: 656  
Score: 1034.00 Matches: 196  
Percent Similarity: 97.60% Conservative: 7  
Best Local Similarity: 94.23% Mismatches: 4  
Query Match: 67.98% Indels: 1  
DB: 14 Gaps: 0  
US-10-042-894A-8 (1-289) x CA245555 (1-656)  
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20  
DB 33 ATGTCGACCTCGCCCGCGGAGCACCAGTCGCGGCCACCGCGCTCCGACACAG 92  
QY 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrIysProLeuGlnAlaGlyAspArg 40  
DB 93 CTGGGTCCGCTCATCGACGCTCTGGCTCTTCTACAAGCGCTCCAGCGTCCGGGACCGC 152  
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
DB 153 GGGAGACAGAGTCGCTCTTACGAGCGCTTCCACCCACCGCCCGCGCGCGCGC 212  
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
DB 213 ATCGGGACACCTTCTCCCGCGGTTCAGGCACGCGACTCTCCACCGGCGCGG 272  
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100  
DB 273 CCAGGGAGCGCATCGCACCTCGTCTCGACGACCTCTCGCGGGCTGGAGGCGCC 332  
QY 101 CysValAlaAspIleIysIleGlyAlaIleThrProSerSerProGluProTyr 120

Db 333 TGCGTCCCGACATCAAGATCGCGGCATCACGTGGCGCGCGAGCTCGCGAGCCCTAC 392  
QY 121 IleAlaAlaCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
DB 393 GTGCCCAAGTGCCTCGCAAGGACCGCGGACCAACGAGCATTCTGCTCGGATTCGCGGTC 452  
QY 141 SerGlyValArgValValGlyProGluGlyValAlaValTrpArgThrGluArgProGluVal 160  
DB 453 TCCGGGTCCGGTCTGTCGTCCTCCGAGGCGCGCTGTGGCGGACTGAGCGCCCGAGGTG 512  
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgValLeuArgArgTyrValSerSerValAla 180  
DB 513 AAGGCGCTGGACACCGCGCGTCCGCGCGCTCTCCGCGCTACGTTTCATCGGTGCC 572  
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLe 200  
DB 573 GACGAGGGGATGAGCTGCGCACTCGCGCGCGGTGTACGCGCGCAAAAGGGGAGTCTT 632  
QY 200 uSerGlnLeuArgGluLeuLys 207  
DB 633 GTCACAGCTGCGGAGCTTAAG 654

RESULT 9  
CA130685  
LOCUS 722 bp mRNA linear EST 24-SEP-2003  
DEFINITION SCCRT1004G05.g RT1 Saccharum officinarum cDNA clone SCCRT1004G05  
5', mRNA sequence.  
CA130685  
ACCESSION CA130685.1 GI:35013914  
VERSION  
KEYWORDS Saccharum officinarum  
SOURCE Saccharum officinarum  
ORGANISM  
REFERENCE Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
AUTHORS The libraries that made SUCEST  
TITLE Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
JOURNAL Contact: Arruda P  
COMMENT Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bcccenter.fcav.unesp.br  
Plate: 004 row: G column: 05  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers  
1..722  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCCRT1004G05"  
/lab\_host="DH10B"  
/clone\_lib="RT1"  
/note="Organ: Root tips (0.3cm-long) from adult plants;  
Vector: pSport1; Site 1: SalI; Site 2: NotI; An  
unidirectional cDNA library generated from [Root tips  
(0.3cm-long) from adult plants]. cDNA was prepared from  
polyA+ mRNA using SuperScript plasmid System Kit  
(Invitrogen). The double-strand cDNAs were fractionated  
in a sepharose CL-2B 40cm-columns and fragments sizing the  
between 0.8 and 1.5 Kb were directionally cloned into the  
vector. Details of each source of RNA and library  
construction can be obtained at  
http://sucest.lad.ic.unicamp.br/public"

ORIGIN  
Alignment Scores:



```

Pred. No.: 1.97e-86 Length: 722
Score: 993.50 Matches: 202
Percent Similarity: 92.07% Conservative: 7
Best Local Similarity: 88.99% Mismatches: 17
Query Match: 65.32% Indels: 4
DB: 13 Gaps: 0

US-10-042-894A-8 (1-289) x CA130685 (1-722)
Qy 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
Db 31 ATGTCGACCTCGC-CGCGCGAGACCAAGTCGCGGCCACCGCGCTCCGCAACAG 89
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 90 CTGGTCCGCTCATCGACGCTCTGGCTCTTCTACAAGCGCTCCAGGTCCGGGACCGC 149
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 150 GGGGAGACGAGCTCGCTCTTACGAGGCGTCTCCACCGCGCGCTCCGCGCCGC 209
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 210 ATCGCGGACACTTCTCCCGCGGTTTCACGGACGCGATCTCTCCACCGAGCGCG 269
Qy 81 ProGlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 270 CCAGGGGAGCGCATCGCACCTCGTCTCGACGACCTCTCGCGGGCTGGAGGCGCC 329
Qy 101 CysValAlaAspLeuLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
Db 330 TGGTCGCGGACATCAAGATCGGCGCATCATCGTGGCGCGGAGCTCGCGGAGCCCTAC 389
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 390 GTGCCCAAGTGCCTCGCCAGGACCGGGGACCAACAGCATCTGCTCGGATTCGGGTC 449
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 450 TCCGGGCTCCGGGCTCGTCCGTCGAGGCGCGCTGTGGGAGACTGAGCGCCGAGGTG 509
Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla 180
Db 510 AAGCCCTGGACACCGCGCGCTCGCGCGCTGCTCGCGCGCTTACGTTATCCGTTGCC 569
Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaAlaValTyrGlyGlyLysGlyGlyValle 200
Db 570 GACGAGGGATGGACTGCGCACTCGCGCGCGGGGACCGCGGCAAGGGGAGTCTT 629
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 630 GTCAAGCTGCCGAGCTTAAAGCGTGAAGCGTGGTTCGAGGA-CAAACTCTGGATCACTT-TACTC 687
Qy 220 rAlaSerIleLeuLeuGly 226
Db 688 GGGGCGCATCTCTTTCGGGC 706

RESULT 10
CG776236 591 bp DNA linear GSS 29-OCT-2003
LOCUS 1123005E08.x1 1123 - RescueMu Grid L Zea mays genomic, genomic
DEFINITION survey sequence.
ACCESSION CG776236
VERSION CG776236.1 GI:38034108
KEYWORDS GSS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 591)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
```

JOURNAL  
COMMENT

Unpublished (2001)  
Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 1123005 column: 3  
Class: transposon-tagged.  
Location/Qualifiers  
I. .591  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="mixed background W23/Al88/B73/K55"  
/db\_xref="taxon:4577"  
/tissue\_type="leaf"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="1123 - RescueMu Grid L"  
/notes="Organ: leaf; Vector: RescueMu (engineered from  
pBluescript backbone); Site 1: BamHI; Site 2: BglII;  
RescueMu is a 4.9 kb, modified maize Mu transposon  
designed to allow plasmid rescue from total genomic DNA.  
Mu elements insert preferentially into transcription  
units. For more information on RescueMu, go to the web  
site 'www.zmdb.iastate.edu' and follow the links for  
'RescueMu.' Grid L was grown in Molokai in 2001. DNA was  
extracted from leaf strips, double digested using BamHI  
and BglII, and ligated to form circular plasmids. DH10B  
cells were transformed and then screened on LB plates with  
ampicillin."

FEATURES  
source

Alignment Scores:  
Pred. No.: 2.03e-86 Length: 591  
Score: 992.00 Matches: 187  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.47% Mismatches: 0  
Query Match: 65.22% Indels: 0  
DB: 29 Gaps: 0

## ORIGIN

US-10-042-894A-8 (1-289) x CG776236 (1-591)  
Qy 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20  
Db 28 ATGTCGACCTCCACCGCGGAGCACCAGTCGCGGCCACCGCGCTCCGCGCAAG 87  
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
Db 88 CTGGGCGCGCTCATCGACGCTCCGCGCTTCTTACAAGCGCTCCAGCGCGGCGCGCT 147  
Qy 41 GlyGluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
Db 148 GGGGAGACGAGGTGCGCTTCTATGAGCGTCTTCGCGCCACCGCGCGCTCCGCGCCGC 207  
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
Db 208 ATCCGAGACACTTCTTCCCGCGGTTCACGCGACGCGACTCTCTCCCAACGAGGCGCAG 267  
Qy 81 ProGlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100  
Db 268 CCGGGGAGCGCATCTCTACCTCGCTCGACGACCTCTCTCGGGGCTTCAGGCGCCC 327  
Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120  
Db 328 TGGTTCGACATCATCAGATCGCGCGCATCATCGTGGCCACCGAGTTCGCGGAGCCCTAC 387  
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 388 ATCGCAAGTGCCTCGCCAAAGGACCGCGGACCAACGAGCGCTTCTGCTCGGATTCGCGTC 447  
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160

Db	448	TCGGGGTCCGAGTCGTCGGCCCGAGGGCGCGTGTGGCGGACGAGCGCCCGAGGTG	507		2	GGGAGCAGAGCTCGCTTCTACGAGCGGTTCTCCAAACCCACGCGCGTGTCCCGGCGCGC	61
QY	161	LysAlaMetAspThrAlaGlyValArgValLeuArgTyrValSerSerValAla	180		61	IIeAtgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln	80
Db	508	AAGGCCATGACACCGCGCGTCCGGCGGTCTCGGGGCTACGTGTCATCCGTGTC	567		62	ATCCGGGACACCTTCTTCCCGCGGTTTCAGGACGCGACTCTCTCCACCGAGGCGCG	121
QY	181	AspGluGlyMetAspCysAlaLeu	188		81	ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro	100
Db	568	GACCAGGGATGGACTGTGGCTC	591		122	CCAGGGAGCGGCATCCGCACCTCGTCTCGACGACTCTCTCGCGGGCTGGAGGGGCC	181
RESULT 11					101	CysValAlaAspIleIleGlyAlaIleThrTrpProSerSerProGluProTyr	120
CA202412					182	TGCGTCCGCGATCATCAAGATCGCGCCATCATCGTGGCCCGGAGCTCGCGGAGCCCTAC	241
LOCUS	CA202412	671 bp	mRNA	linear	121	IIeAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal	140
DEFINITION	SCRLFL1009H02.9 Fl1 Saccharum officinarum cDNA clone	SCRLFL1009H02			242	GTCGCCAAGTGCTCGCCAAAGCACCGCGGACCCAGGACATTCGTCTCGGATTCGGGTC	301
ACCESSION	CA202412				141	SerGlyValArgValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal	160
VERSION	CA202412.1	GI:35237651			302	TCCGGCGTCCGGTCTCGTCCGAGGCGCGTGTGGCGGACTGAGCGCCGAGGTG	361
KEYWORDS	EST.				161	LysAlaMetAspThrAlaGlyValArgValLeuArgTyrValSerSerValAla	180
SOURCE	Saccharum officinarum				362	AAGGCCCTGGACACCGCGCGTCCGCGCGTCTCGCGGCTACGTTTCATCCGTGTC	421
ORGANISM	Saccharum officinarum				181	AspGluGlyMetAspCysAlaLeuAlaValTyrGlyLysGlyGlyValLeu	200
REFERENCE	Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.				422	GACGAGGATGAGCTGCGCACTCCCGCGCGGTGTACGGCGGAGGAGGTCTTG	481
AUTHORS	The libraries that made SUCST				201	SerGlnLeuArgGluLeuLysAlaTyrPheGluGluGlnThrLeuPheHisPheTyrSer	220
TITLE	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)				482	TCACAGCTCGCGAGCTTAAGCGTGGTTCGAGGAGCAGACTCTGCTCCACTTCTACTCG	541
JOURNAL	Contact: Arruda P				221	AlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyGlyGlyGly	240
COMMENT	Centro de Biologia Molecular e Engenharia Genetica				542	CGCGTGAATCTTTTGGGCTATGATGCTAGTCAAGTCAGCAGCGGAGGAGGTGAAGTGGG	601
	Universidade Estadual de Campinas				241	ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp-GlyValIleAspHi	260
	Caixa Postal 6010, 13083-970, Campinas SP, Brazil				602	TTGAAGGTGAA-CTTGTTCACCTTTCCCAATTTTCCCAAGTTGATAGGGGGAATGACCA	660
	Tel: 55 19 3788 1137				260	sAsnPhe	262
	Fax: 55 19 3788 1089				661	AAATTC	667
	Email: parruda@unicamp.br						
	Clone distribution: clone distribution information can be found						
	through the Brazilian Clone Collection Center (BCCC) at						
	http://www.bccc.org.br						
	Plate: 009 row: H column: 02						
	Seq primer: T7 Promoter Primer.						
FEATURES	Location/Qualifiers						
source	1..671						
	/organism="Saccharum officinarum"						
	/mol_type="mRNA"						
	/db_xref="taxon:4547"						
	/clone="SCRLFL1009H02"						
	/lab_host="DH10B"						
	/clone_lib="FL1"						
	/note="Organ: Inflorescence at beginning of development						
	(1cm-long); Vector: pSport1; Site_1: SalI; Site_2: NotI;						
	An unidirectional cDNA library generated from						
	[inflorescence at beginning of development (1cm-long)].						
	cDNA was prepared from polyA+ mRNA using SuperScript						
	Plasmid System Kit (Invitrogen). The double-strand cDNAs						
	were fractionated in a sepharose CL-2B 40cm-columns and						
	fragments sizing between 0.8 and 1.5 Kb were						
	directionally cloned into the vector. Details of each						
	source of RNA and library construction can be obtained at						
	http://sucst.lad.ic.unicamp.br/public"						
ORIGIN							
Alignment Scores:							
Pred. No.:	7.58e-86	Length:	671				
Score:	987.00	Matches:	192				
Percent Similarity:	91.93%	Conservative:	13				
Best Local Similarity:	86.10%	Mismatches:	17				
Query Match:	64.89%	Indels:	2				
DB:	14	Gaps:	0				
US-10-042-894A-8 (1-289) x CA202412 (1-671)							
QY	41	GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg	60				

these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).

## FEATURES

source  
 Location/Qualifiers  
 1..3374  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="maizdb:632179"  
 /clone\_lib="Maize Mapping Project/DuPont Cornsensus Library"  
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

## ORIGIN

Alignment Scores:  
 Pred. No.: 3.45e-84 Length: 3374  
 Score: 981.00 Matches: 205  
 Percent Similarity: 71.88% Conservative: 2  
 Best Local Similarity: 71.18% Mismatches: 62  
 Query Match: 64.50% Indels: 20  
 DB: 11 Gaps: 1

US-10-042-894A-8 (1-289) x AY109355 (1-3374)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
 DB 72 ATGCCGACTTCCACCCGCGGAGACCAAGTCGCGGTCCACCGGCTCGCCAGCAAG 131  
 QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
 DB 132 CTGGGCCACTCATCGACGACTCTGGCTCTTCAAGCGCTCCACGCGCGGCGACCGT 191  
 QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
 DB 192 GGGGAGCAGGAGGTCGCTTCTATGAGCGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 251  
 QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
 DB 252 NTCCGAGACACTTCTNNNNNGTTCCACGGCAGCGACTCTCCACCGAGGCGCAG 311  
 QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100  
 DB 312 CCGCGGAGCGCATCCGCACTCTCTCGACGACCTCTCGCGGGGTTTGAGCGGCC 371  
 QY 101 CysValAlaAspIleLysIleGlyAlaIlePheThrProProSerSerProGluProTyr 120  
 DB 372 TGCCTCGAGACATCAAGATCGGTGCCATCACGTG----- 406  
 QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
 DB 407 -----ACCACGAGCGTTCGTCTCGGATTCGCGTCC 436  
 QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
 DB 437 TCCGCGCTCCGAGTCGTCCGCGCGGCGCGCGTGNNNNNNNNNNNNNNNNNNNNN 496  
 QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAl 180  
 DB 497 AAGGCTATGAGCATTTGTCGCGCTCCGCGCTGTCTCGCGCTACGTGTATCCGCTTGC 556  
 QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLe 200  
 DB 557 CGACAGGGGATGAGCTCGCGCTCGCGCGCGGCTGTACGAGAGAAAGGTGAGTCTT 616  
 QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220  
 DB 617 GTACAGCTCGCGAGCTCAAGCGTGTGTTTCGAGGGGCAGACTNNNNNNNNNNNNNN 676

QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyGlyGlyGly 240  
 DB 677 NNN 736  
 QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValLleAspHi 260  
 DB 737 NNN 796  
 QY 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280  
 DB 797 CAACCTTCCTGGCGGGCTCTGCTAGCTGATCAAGTTTGTTCATGACATTTTCCAGAGAC 856  
 QY 280 rProHisThrGlnProLeuGly 287  
 DB 857 TCCTCAGACGACCCCTTTGGN 878

RESULT 13  
 CC724950/c  
 LOCUS  
 DEFINITION  
 CC724950 722 bp DNA linear GSS 23-JUN-2003  
 OGLAB23TH ZM 0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0307C22,  
 genomic survey sequence.  
 CC724950  
 VERSION  
 CC724950.1 GI:32143883  
 KEYWORDS  
 GSS.  
 SOURCE  
 Zea mays  
 ORGANISM  
 Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 722)  
 Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
 Citek, R.W., Nummer, A., Robbins, D. and Lakey, N.  
 Consortium for Maize Genomics  
 Unpublished (2002)  
 Other GSSs: OGLAB23TV  
 Contact: Cathy Whitelaw  
 TIGR Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TR  
 Class: sheared ends.  
 Location/Qualifiers  
 1..722  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBMA0307C22"  
 /clone\_lib="ZM\_0.7\_1.5\_KB"  
 /notes="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

FEATURES  
 source  
 1..722  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBMA0307C22"  
 /clone\_lib="ZM\_0.7\_1.5\_KB"  
 /notes="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1e-84 Length: 722  
 Score: 976.00 Matches: 197  
 Percent Similarity: 88.44% Conservative: 2  
 Best Local Similarity: 87.56% Mismatches: 7  
 Query Match: 64.17% Indels: 20  
 DB: 29 Gaps: 1

US-10-042-894A-8 (1-289) x CC724950 (1-722)

QY 66 PheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGluProHis 85  
 DB 721 TTCCCCCGGTTCACGGCAGCGACTCTCCCCACCGAGCGCAGCCCGGGGCGCGCAT 662  
 QY 86 ProHisLeuValLeuAspLeuAlaGlyPheGlnAlaProCysValAlaAspIle 105  
 DB 661 CCGCACCTGCTCTCGACGACCTCTCGCGGGGTTTGAGGCGGCTGCGTCCGAGACATC 602

```
QY 106 LysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAlaLysCysLeu 125
Db 601 AAGATCGGTGCCATCACGTG-----582
QY 126 AlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgVal 145
Db 581 -----ACCAAGAGCGTTCGCTCGGATTCGGGTCTCCGGGCTCCGAGTC 537
QY 146 ValGlyProGluGlyAlaValTrpArgThrGluArgProGluValLysAlaMetAspThr 165
Db 536 GTCGGCCCGAGGGCGCGTGTGGCGAGCGAGCGCCGAGGTGAAGCTATGACATT 477
QY 166 AlaGlyValArgValLeuArgArgTyrValSerSer-ValAlaAspGluGlyMetAs 185
Db 476 GTCGGCGTCCGGCGGTGTCGGCGCTACGTGTATCCGCTTGCAGAGAGGGATGGA 417
QY 185 pCysAlaLeuAlaAlaValTyrGlyLysGlyValLeuSerGlnLeuArgG 205
Db 416 CTGGCGCTCGCGCGCGCGGTACGGAGGAAGGTGGAGTCTGTACAGCTGGCGGA 357
QY 205 uLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSerAlaSerIleLeu 225
Db 356 GCTCAGGCGTGTTCGAGGGCGAGCTCTGTCCACTTCTACTCGGCTCGATTCTTCT 297
QY 225 uGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyValThrValLysLe 245
Db 296 GGGCTATGATGCTGTCTGCTGAGTCGACAGCGGAGGTGGGGTGAACAGTGAAGCT 237
QY 245 uValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHisAsnPheLeuGly 265
Db 236 GTTGAGACTTTCGCCAGTGGCGAGGTGATGGGGTATGACCAACTTCTCTGGGGG 177
QY 265 yLeuCysserLeuIleLysPheValSerAspIleValProGluThrProHisThrGlnPr 285
Db 176 GCTCTGTAGCTGATCAAGTTTGTCTGACATTGTTCCAGAGACTCTCTCAGAGCAGCC 117
QY 285 oLeuGlyProSer 289
Db 116 TTGGGTCTCTCT 104

RESULT 14
CAL130686
LOCUS
DEFINITION
5', mRNA sequence.
EST 24-SEP-2003
SCCCT1004G06.g RT1 Saccharum officinarum cDNA clone SCCCT1004G06
CAL130686
VERSION
CAL130686.1 GI:35013916
KEYWORDS
EST.
SOURCE
Saccharum officinarum
ORGANISM
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.
1 (bases 1 to 744)
Vettore,A.L., da Silva,P.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 004 row: G column: 06
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..744
/organism="Saccharum officinarum"
```

```
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCCT1004G06"
/lab_host="DH10B"
/clone_lib="RT1"
/note="Organ: Root tips (0.3cm-long) from adult plants;
Vector: pSport1; Site_1: SalI; Site_2: NotI; An
unidirectional cDNA library generated from [Root tips
(0.3cm-long) from adult plants]. cDNA was prepared from
polyA+ mRNA using Superscript Plasmid System Kit
[Invitrogen]. The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucst.lad.ic.unicamp.br/public"
```

## ORIGIN

```
Alignment Scores:
Pred. No.: 2,06e-80 Length: 744
Score: 932.00 Matches: 191
Percent Similarity: 86.02% Conservative: 12
Best Local Similarity: 80.93% Mismatches: 26
Query Match: 61.28% Indels: 9
DB: 13 Gaps: 1

US-10-042-894A-8 (1-289) x CAL130686 (1-744)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 31 ATGTCGGACCTCGC-CCGCGGAGAGACCAAGTCGCGGCCACCGCGCTCCGCCAACAG 89
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 90 CTGGTCCGCTCATCGACGCTCTGGCTCTTCTACAAAGCCGCTCAGGTCGGGACCGC 149
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 150 GGGAGACGAGCTCGCTTCTACAGGCGTTCACCCAGCGCGCTCCGCCCGCC 209
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 210 ATCCGGGACACCTTCTTCCCGCGGTTTCAGGCGACGCGACTCTCTCCACCGAGCGCG 269
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 270 CCAGGGAGCGCATCCGCACTCTCTCTCGACGACCTCTCTCGCGGGCTCGAGGCGGCC 329
QY 101 CysValAlaAspIleLysIleGlyValAlaIleThrTrpProProSerSerProGluProTyr 120
Db 330 TCGTCCGCGACATCAAGATCGCGCCATCATCGTCCGCGCGAGCTCGCGGAGCGCTTAC 389
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 390 GTCGCCAAGTCTCTGCCAAGGACCGCGGACCCAGCAGCATTCGTCGGATTCCGCGTC 449
QY 141 SerGlyValArgValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 450 TCCGGCTCCGGGTGCTGCTCCGAGGGCGCGCTGGGGCGGACTGAAACGCCCGAGGGT 509
QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgTyrValSerSerValAla 180
Db 510 AAGGCCCTGGACACCGCGCGCTCCGCCGGGGCTTCGCGCGCTAGCTTTATCTGTGTC 569
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLeu 200
Db 570 GACGAGGGGATGACTGCTCCCTCCCGCGCGGT-TACGGGGGCAAAAGGGGAACTTG 628
QY 201 SerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSer 220
Db 629 GTAACATGGGCAAGCTTAAAGGGGGGTAGAGGGGAAAAATTTGGTTC----- 676
QY 221 AlaSerIleLeu-LeuGlyTyrAspAlaAlaValAlaAlaGly 235
```

Db	677	-----CTTTTAATGGGGGTAGATCTTTGGGGTATTATTCAGGG	716
RESULT 15			
CK211413/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

Search completed: March 27, 2004, 08:26:37  
Job time : 3179 secs

## ORIGIN

Alignment Scores:

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 20:51:24 ; Search time 60 Seconds  
(without alignments)  
1360.937 Million cell updates/sec

Title: US-10-042-894a-8

Perfect score: 1521

Sequence: 1 MSLLHPPEHQVAGHRASAK.....IKFVSIVPETHQTPLGPS 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Genesep 29Jan04:\*

1: Genesep1980s:\*

2: Genesep1990s:\*

3: Genesep2000s:\*

4: Genesep2001s:\*

5: Genesep2002s:\*

6: Genesep2003as:\*

7: Genesep2003bs:\*

8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1521	100.0	289	5	Aae26196 Maize ino
2	1485	97.6	289	5	Aae26195 Maize ino
3	914.5	60.1	227	5	Aae26194 Maize ino
4	748.5	49.2	240	5	Aae26193 Maize ino
5	742.5	48.8	295	5	Aae26199 P. argent
6	722.5	47.5	310	5	Aae26198 Eucalyptu
7	714.5	47.0	300	3	Aag46591 Arabidops
8	714.5	47.0	330	3	Aag46590 Arabidops
9	700.5	46.1	300	3	Aag44257 Arabidops
10	700.5	46.1	330	3	Aag44256 Arabidops
11	674.5	44.3	286	3	Aag19913 Arabidops
12	674.5	44.3	286	3	Aag20765 Arabidops
13	674.5	44.3	286	3	Aag06976 Arabidops
14	674.5	44.3	327	3	Aag06975 Arabidops
15	673.5	44.3	286	5	AB933390 Herbicida
16	644.5	42.4	279	5	Aae26197 Soybean i
17	578	38.0	111	5	Aae26200 Maize ino
18	577	37.9	111	5	Aae26201 Maize ino
19	542.5	35.7	248	3	Aag46592 Arabidops
20	494	32.5	205	3	Aag19915 Arabidops
21	494	32.5	205	3	Aag20767 Arabidops
22	494	32.5	206	3	Aag20766 Arabidops
23	494	32.5	206	3	Aag19914 Arabidops
24	494	32.5	206	3	Aag06977 Arabidops
25	439.5	28.9	202	3	Aag24258 Arabidops

26	265	17.4	291	4	ABE68470	Abb68470 Drosophil
27	201.5	13.2	416	6	AAE34863	Aae34863 Human kin
28	192	12.6	410	5	AAE24144	Aae24144 Human kin
29	192	12.6	410	7	ADC06743	Adc06743 Human ino
30	191	12.6	293	6	ADA55092	Ada55092 Human pro
31	183	12.0	353	4	AA41451	Aam41451 Human pol
32	183	12.0	426	4	AA39665	Aam39665 Human pol
33	183	12.0	426	4	AA394867	Aab94867 Human pro
34	183	12.0	426	4	AAG89197	Aag89197 Human sec
35	183	12.0	426	4	AAG89190	Aag89190 Human sec
36	183	12.0	426	7	ADC06747	Adc06747 Human ino
37	183	12.0	481	4	AAG89295	Aag89295 Human sec
38	182.5	12.0	396	7	ADC06749	Adc06749 Murine in
39	177	11.6	426	3	AA42986	Aab42986 Human ORF
40	163.5	10.7	441	7	ADC06745	Adc06745 Human ino
41	147	9.7	477	5	ABB09782	Abb09782 Amino aci
42	145	9.5	461	5	ABB09772	Abb09772 Amino aci
43	142.5	9.4	441	5	ABB09770	Abb09770 Amino aci
44	142.5	9.4	441	5	ABB09780	Abb09780 Amino aci
45	139	9.1	313	4	AAO7085	Aao7085 Human pol

## ALIGNMENTS

## RESULT 1

AAE26196

ID AAE26196 standard; protein; 289 AA.

XX

AC AAE26196;

XX

DT 14-NOV-2002 (first entry)

XX

DE Maize inositol polyphosphate kinase (IPPK) #4.

XX

KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;

KW nutritional value; animal feed; transgenic.

XX

OS Zea mays.

XX

PN WO200259324-A2.

XX

PD 01-AUG-2002.

XX

PF 09-JAN-2002; 2002WO-US003120.

XX

PR 12-JAN-2001; 2001US-0261465P.

XX

(PION-) PIONEER HI-BRED INT INC.

XX

PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;

 XX || DR | WPI; 2002-636540/68. |
DR	N-PSDB; AAD43514.
XX	
PT	New inositol polyphosphate kinase polynucleotides and polypeptides,
PT	useful in modulating phytic acid biosynthesis by decreasing phytate or
PT	increasing non-phytate phosphorous to improve the nutritional value of
PT	animal feed.
XX	
PS	Claim 12; Page 65-66; 86pp; English.
XX	
CC	The invention relates to novel inositol polyphosphate kinase (IPPK)
CC	polypeptides and polynucleotides. Sequences of the invention are useful
CC	in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC	increasing non-phytate phosphorous to improve the nutritional value of
CC	animal feed, or to reduce the environmental impact of animal waste.
CC	Polynucleotides of the invention are to produce transgenic plants with an
CC	altered phenotype. IPPK proteins are used to screen compounds that
CC	modulate their activity and raising anti-idiotypic antibodies. The
CC	present sequence is maize IPPK protein
XX	
SQ	Sequence 289 AA;

Query Match 100.0%; Score 1521; DB 5; Length 289;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-159;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDLHPPEHOVAGHRASAKGLPLIDSGGLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60  
 |||||  
 Db 1 MSDLHPPEHOVAGHRASAKGLPLIDSGGLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60  
 |||||

QY 61 IRDTFFPRFHGTRLLPTEAQGPCEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120  
 |||||  
 Db 61 IRDTFFPRFHGTRLLPTEAQGPCEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120  
 |||||

QY 121 IAKCLAKDRGTTSVLLGFRVSGVRVVGPEGAVWRTERPEVKAMDTAGVRRVLRYYSSVA 180  
 |||||  
 Db 121 IAKCLAKDRGTTSVLLGFRVSGVRVVGPEGAVWRTERPEVKAMDTAGVRRVLRYYSSVA 180  
 |||||

QY 181 DEGMDCALAAAVYGGKGVLSQLRELKAWFEETLHFYFYSASILLGYDAAAAGGDDGG 240  
 |||||  
 Db 181 DEGMDCALAAAVYGGKGVLSQLRELKAWFEETLHFYFYSASILLGYDAAAAGGDDGG 240  
 |||||

QY 241 VTKLVDFAHVABGDGVIDHNFGLGGLCSLIKFFVSDIVPETHPTQPLGPS 289  
 |||||  
 Db 241 VTKLVDFAHVABGDGVIDHNFGLGGLCSLIKFFVSDIVPETHPTQPLGPS 289  
 |||||

RESULT 2  
 AAE26195  
 ID AAE26195 standard; protein; 289 AA.  
 XX  
 AC AAE26195;  
 XX  
 DT 14-NOV-2002 (first entry)  
 XX  
 DE Maize inositol polyphosphate kinase (IPPK) #3.  
 XX  
 KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
 KW nutritional value; animal feed; transgenic.  
 XX  
 OS Zea mays.  
 XX  
 PN WO200259324-A2.  
 XX  
 PD 01-AUG-2002.  
 XX  
 PF 09-JAN-2002; 2002WO-US003120.  
 XX  
 PR 12-JAN-2001; 2001US-0261465P.  
 XX  
 PA (PION-) PIONEER HI-BRED INT INC.  
 XX  
 PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
 XX  
 DR WPI; 2002-636540/68.  
 DR N-PSDB; AAD43513.  
 XX  
 PT New inositol polyphosphate kinase polynucleotides and polypeptides,  
 PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
 PT increasing non-phytate phosphorous to improve the nutritional value of  
 PT animal feed.  
 XX  
 PS Claim 12; Page 63-64; 86pp; English.  
 XX  
 CC The invention relates to novel inositol polyphosphate kinase (IPPK)  
 CC polypeptides and polynucleotides. Sequences of the invention are useful  
 CC in modulating the phytic acid biosynthesis by decreasing phytate and/or  
 CC increasing non-phytate phosphorous to improve the nutritional value of  
 CC animal feed, or to reduce the environmental impact of animal waste.  
 CC Polynucleotides of the invention are used to produce transgenic plants with an  
 CC altered phenotype. IPPK proteins are used to screen compounds that  
 CC modulate their activity and raising anti-idiotypic antibodies. The  
 CC present sequence is maize IPPK protein

SQ Sequence 289 AA;  
 Query Match 97.6%; Score 1485; DB 5; Length 289;  
 Best Local Similarity 97.9%; Pred. No. 5e-155;  
 Matches 283; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSDLHPPEHOVAGHRASAKGLPLIDSGGLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60  
 |||||  
 Db 1 MPDLHPPEHOVAGHRASAKGLPLIDSGGLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60  
 |||||

QY 61 IRDTFFPRFHGTRLLPTEAQGPCEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120  
 |||||  
 Db 61 IRDTFFPRFHGTRLLPTEAQGPCEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120  
 |||||

QY 121 IAKCLAKDRGTTSVLLGFRVSGVRVVGPEGAVWRTERPEVKAMDTAGVRRVLRYYSSVA 180  
 |||||  
 Db 121 IAKCLAKDRGTTSVLLGFRVSGVRVVGPEGAVWRTERPEVKAMDTAGVRRVLRYYSSVA 180  
 |||||

QY 181 DEGMDCALAAAVYGGKGVLSQLRELKAWFEETLHFYFYSASILLGYDAAAAGGDDGG 240  
 |||||  
 Db 181 DEGMDCALAAAVYGGKGVLSQLRELKAWFEETLHFYFYSASILLGYDAAAAGGDDGG 240  
 |||||

QY 241 VTKLVDFAHVABGDGVIDHNFGLGGLCSLIKFFVSDIVPETHPTQPLGPS 289  
 |||||  
 Db 241 VTKLVDFAHVABGDGVIDHNFGLGGLCSLIKFFVSDIVPETHPTQPLGPS 289  
 |||||

RESULT 3  
 AAE26194  
 ID AAE26194 standard; protein; 227 AA.  
 XX  
 AC AAE26194;  
 XX  
 DT 14-NOV-2002 (first entry)  
 XX  
 DE Maize inositol polyphosphate kinase (IPPK) #2.  
 XX  
 KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
 KW nutritional value; animal feed; transgenic.  
 XX  
 OS Zea mays.  
 XX  
 PN WO200259324-A2.  
 XX  
 PD 01-AUG-2002.  
 XX  
 PF 09-JAN-2002; 2002WO-US003120.  
 XX  
 PR 12-JAN-2001; 2001US-0261465P.  
 XX  
 PA (PION-) PIONEER HI-BRED INT INC.  
 XX  
 PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
 XX  
 DR WPI; 2002-636540/68.  
 DR N-PSDB; AAD43512.  
 XX  
 PT New inositol polyphosphate kinase polynucleotides and polypeptides,  
 PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
 PT increasing non-phytate phosphorous to improve the nutritional value of  
 PT animal feed.  
 XX  
 PS Claim 12; Page 61-62; 86pp; English.  
 XX  
 CC The invention relates to novel inositol polyphosphate kinase (IPPK)  
 CC polypeptides and polynucleotides. Sequences of the invention are useful  
 CC in modulating the phytic acid biosynthesis by decreasing phytate and/or  
 CC increasing non-phytate phosphorous to improve the nutritional value of  
 CC animal feed, or to reduce the environmental impact of animal waste.  
 CC Polynucleotides of the invention are used to produce transgenic plants with an  
 CC altered phenotype. IPPK proteins are used to screen compounds that  
 CC modulate their activity and raising anti-idiotypic antibodies. The  
 CC present sequence is maize IPPK protein



```

XX SQ Sequence 227 AA;
Query Match 60.1%; Score 914.5; DB 5; Length 227;
Best Local Similarity 78.5%; Pred. No. 3.8e-92;
Matches 183; Conservative 8; Mismatches 35; Indels 7; Gaps 2;

QY 1 MSDLHPPEHQVAGHRASASKGLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAPAR 60
Db 1 MPDLHPPEHQVAGHRASASKGLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAPAR 60
QY 61 IRDTFFPRFHGTRLLPTEAQGPGEHPHLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120
Db 61 IRDTFFPRFHGTRLLPTEAQGPGEHPHLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120
QY 121 IAKCLAKDRGTTSVLLGFRVVGPEGAVWTERPEVKAMDTAGVRRLRYVSSVA 180
Db 121 IAKCLAKDRGTTSVLLGFRVVGPEGAVWTERPEVKAMDTAGVRRLRYVSSVA 180
QY 181 DEGMDCALAAAVYGGKGVLSQRLKAWFEEQTLPH-FYSASILLGYDAAAV 232
Db 180 -----CRGDGLRARGGVRRKWSLVTARAQGVVGAASVPLLIGVDSSGL 227

RESULT 4
AAE26193
ID AAE26193 standard; protein; 240 AA.
AC AAE26193;
XX
XX DT 14-NOV-2002 (first entry)
XX DE Maize inositol polyphosphate kinase (IPPK) #1.
XX KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
XX KW nutritional value; animal feed; transgenic.
XX OS Zea mays.
XX PN WO200259324-A2.
XX PD 01-AUG-2002.
XX PF 09-JAN-2002; 2002WO-US003120.
XX PR 12-JAN-2001; 2001US-0261465P.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX DR N-PSDB; AAD43511.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX PT useful in modulating phytic acid biosynthesis by decreasing phytate or
XX PT increasing non-phytate phosphorous to improve the nutritional value of
XX PT animal feed.
XX PS Claim 12; Page 59-60; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX CC polypeptides and polynucleotides. Sequences of the invention are useful
XX CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
XX CC increasing non-phytate phosphorous to improve the nutritional value of
XX CC animal feed, or to reduce the environmental impact of animal waste.
XX CC Polynucleotides of the invention are to produce transgenic plants with an
XX CC altered phenotype. IPPK proteins are used to screen compounds that
XX CC modulate their activity and raising anti-idiotypic antibodies. The
XX CC present sequence is maize IPPK protein
XX SQ Sequence 240 AA;

Query Match 49.2%; Score 748.5; DB 5; Length 240;
Best Local Similarity 64.7%; Pred. No. 9.1e-74;
Matches 163; Conservative 10; Mismatches 50; Indels 29; Gaps 6;

QY 1 MSDLHPPEHQVAGHRASASKGLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAPAR 60
Db 1 MPDLHPPEHQVAGHRASASKGLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAPAR 60
QY 61 IRDTFFPRFHGTRLLPTEAQGPGEHPHLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120
Db 61 IRDTFFPRFHGTRLLPTEAQGPGEHPHLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120
QY 121 IAKCLAKDRGTTSVLLGFRVVGPEGAVWTERPEVKAMDTAGVRRLRYVSSVA 164
Db 121 IAKCLAKDRGTTSVLLGFRVVGPEGAVWTERPEVKAMDTAGVRRLRYVSSVA 180
QY 165 TAG-VRRVLRYYSSVADGMDCALAAAVYGGKGVLSQRLKAWFEEQTLPHFYSASI 223
Db 181 TRGWTARSRRCCTEEKVESHSC--ASSRHGWSRLCS-----TSTRRRFFWAMM 228

RESULT 5
AAE26199
ID AAE26199 standard; protein; 295 AA.
XX
XX AC AAE26199;
XX DT 14-NOV-2002 (first entry)
XX DE P. argentatum inositol polyphosphate kinase (IPPK) #1.
XX KW Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
XX KW nutritional value; animal feed.
XX OS Parthenium argentatum.
XX PN WO200259324-A2.
XX PD 01-AUG-2002.
XX PF 09-JAN-2002; 2002WO-US003120.
XX PR 12-JAN-2001; 2001US-0261465P.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX DR N-PSDB; AAD43517.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX PT useful in modulating phytic acid biosynthesis by decreasing phytate or
XX PT increasing non-phytate phosphorous to improve the nutritional value of
XX PT animal feed.
XX PS Claim 12; Page 72-73; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX CC polypeptides and polynucleotides. Sequences of the invention are useful
XX CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
XX CC increasing non-phytate phosphorous to improve the nutritional value of
XX CC animal feed, or to reduce the environmental impact of animal waste.
XX CC Polynucleotides of the invention are to produce transgenic plants with an
XX CC altered phenotype. IPPK proteins are used to screen compounds that
XX CC modulate their activity and raising anti-idiotypic antibodies. The
XX CC present sequence is P. argentatum IPPK protein
XX SQ Sequence 295 AA;

```

```

Query Match      48.8%; Score 742.5; DB 5; Length 295;
Best Local Similarity 52.1%; Pred. No. 5.7e-73;
Matches 149; Conservative 45; Mismatches 79; Indels 13; Gaps 5;

QY 4 LHPPEHQVAGHRASAKLGLPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPARIRD 63
DB 2 LKAPDQVAGHEAGLGLPLIDGSLFYKPLQDNGSEVAFYEFSSNNIPEHIR- 60
QY 64 TFFPRFHGTRLLPTEAQFGEPPHPLVLDLLAGFAQPCVADIKIGAITWPPSSPEPIAK 123
DB 61 KFFPIYVGTKM--KASTGSDHPHMLVLDLTSAHVNSVMDIKIGSRTPAEPASEAYIAK 118
QY 124 CLAKDRGTTSVLLGFRVSGVRVGPGEA-VWRTERPEVKAMDTAGVRRVLRVYSSVADE 182
DB 119 CLKKDRGTTSVLLGFRVSGVRVGPGEA-VWRTERPEVKAMDTAGVRRVLRVYSSVADE 178
QY 183 -----GMDCALAAAVYGGKGVLSQRLKAWFEEQTLPHFYASILLGYDAAVA 234
DB 179 EMWRTGIGPDCSLASFYGGPGLIQLMELKTFWEDQTIYHYACSFLEFIEKRLVLK 238
QY 235 GGDGGVTVKLVDFAHVAAGGVLDHNFGLGCLSLIKFVSDIVPET 280
DB 239 GA-RSNAEVLIDFAHVTGNGVIDHNFGLGCLSLIKFISDILSET 283

RESULT 6
AAE26198
ID AAE26198 standard; protein; 310 AA.
XX AC AAE26198;
XX DT 14-NOV-2002 (first entry)
XX DE Eucalyptus grandis inositol polyphosphate kinase (IPPK).
XX KW Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
XX KW nutritional value; animal feed.
XX OS Eucalyptus grandis.
XX PN WO200259324-A2.
XX PD 01-AUG-2002.
XX PF 09-JAN-2002; 2002WO-US003120.
XX PR 12-JAN-2001; 2001US-0261465P.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX DR N-PSDB; AAD43516.
XX PT New inositol polyphosphate kinase polynucleotides and polypeptides,
XX PT useful in modulating phytic acid biosynthesis by decreasing phytate or
XX PT increasing non-phytate phosphorous to improve the nutritional value of
XX PT animal feed.
XX PS Claim 12; Page 70-71; 86pp; English.
XX CC The invention relates to novel inositol polyphosphate kinase (IPPK)
XX CC polypeptides and polynucleotides. Sequences of the invention are useful
XX CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
XX CC increasing non-phytate phosphorous to improve the nutritional value of
XX CC animal feed, or to reduce the environmental impact of animal waste.
XX CC Polynucleotides of the invention are to produce transgenic plants with an
XX CC altered phenotype. IPPK proteins are used to screen compounds that
XX CC modulate their activity and raising anti-idiotypic antibodies. The
XX CC present sequence is Eucalyptus grandis IPPK protein

SQ Sequence 310 AA;
Query Match      47.5%; Score 722.5; DB 5; Length 310;
Best Local Similarity 52.8%; Pred. No. 9.9e-71;
Matches 149; Conservative 38; Mismatches 84; Indels 11; Gaps 6;

QY 4 LHPPEHQVAGHRASAKLGLPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPARIRD 63
DB 2 LKVPDQVAGHRGSGKGLPLVDGSRFYKPLQSDHRGDETEVAFYFYSNTLIPGHIR- 60
QY 64 TFFPRFHGTRLLPTEAQFGEPPHPLVLDLLAGFAQPCVADIKIGAITWPPSSPEPIA 122
DB 61 KFFPAFHGTKTI--EASDGSQPQPHLVLEDLVSGRTNPSLMDIKTSGRTWYPEASEEYIQ 118
QY 123 KCLAKDRGTTSVLLGFRVSGVRV-VPEGAVWTERPEVKAMDTAGVRRVLRVYSSVAD 181
DB 119 KCLEKDRNSTSVSLGFRISGLRVYQNSEAGFWQPEKKVYVSFNADGVRSLRKPFVSSNLS 178
QY 182 EG---MDCALAAAVYGGKGVLSQRLKAWFEEQTLPHFYASILLGYDAAVAAGGD 237
DB 179 LGNVDPDCLYASKVYCHRGGLIQLQLKENFEVQTNHYFYSCSLILLYDRESALDG-- 236
QY 238 GGGVTVKLVDFAHVAAGGVLDHNFGLGCLSLIKFVSDIVPE 279
DB 237 CAHPKVKLVDFAHVMDGHGVLDHNFGLGCLSVIKFIRDIAD 278

RESULT 7
AAG46591
ID AAG46591 standard; protein; 300 AA.
XX AC AAG46591;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 58631.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 23-APR-1999; 99US-0130891P.
XX PR 28-APR-1999; 99US-0131443P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132407P.
XX PR 05-MAY-1999; 99US-0132484P.
XX PR 06-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 07-MAY-1999; 99US-0132487P.
XX PR 11-MAY-1999; 99US-0132863P.
XX PR 14-MAY-1999; 99US-0134218P.
```

PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 07-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 10-JUN-1999; 99US-0139119P.  
PR 14-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 24-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148315P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.

PR	25-OCT-1999;	99US-0161404P.	PR	30-APR-1999;	99US-0132048P.
PR	25-OCT-1999;	99US-0161405P.	PR	30-APR-1999;	99US-0132407P.
PR	25-OCT-1999;	99US-0161406P.	PR	04-MAY-1999;	99US-0132484P.
PR	26-OCT-1999;	99US-0161359P.	PR	05-MAY-1999;	99US-0132485P.
PR	26-OCT-1999;	99US-0161360P.	PR	06-MAY-1999;	99US-0132486P.
PR	26-OCT-1999;	99US-0161361P.	PR	07-MAY-1999;	99US-0132487P.
PR	28-OCT-1999;	99US-0161920P.	PR	11-MAY-1999;	99US-0132863P.
PR	28-OCT-1999;	99US-0161922P.	PR	14-MAY-1999;	99US-0134256P.
PR	28-OCT-1999;	99US-0161933P.	PR	14-MAY-1999;	99US-0134218P.
PR	29-OCT-1999;	99US-01621142P.	PR	14-MAY-1999;	99US-0134219P.
Query Match 47.0%; Score 714.5; DB 3; Length 300; Best Local Similarity 52.8%; Pred. No. 7.2e-70; Matches 150; Conservative 39; Mismatches 82; Indels 13; Gaps 8;					
Qy	4	LHPPEHVAGHRAASKLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAVPAIRID 63	PR	20-MAY-1999;	99US-0135124P.
Db	2	LKVPEHVAGHIAIDGKGLPLVDQGRFFKPLQGRSHEAKFYESFTSNMKVPDIH- 60	PR	21-MAY-1999;	99US-0135353P.
Qy	64	TFPRFHGTRLTPTEAQGPBH-PHLVLDLLAGFQPCVADIKIGALTWPPSPPEYIA 122	PR	24-MAY-1999;	99US-0135629P.
Db	61	RYFPVYHGTQIV--EASDGSGKPLHLVLDVVGSGYANPSVMDVKIGRTWYPDVSEYFK 118	PR	25-MAY-1999;	99US-0136021P.
Qy	123	KCLAKDRGTTSVILGFRVSGVRVVG-PEGAVWRTERPEVKAMDTAGVRRVLRXYVS--SV 179	PR	27-MAY-1999;	99US-0136392P.
Db	119	KCIKKDRQTTVSLGFRVSGFKIFDHQESSFWRAEKKLVLGYNADGARLALRKFSVNSNP 178	PR	28-MAY-1999;	99US-0136782P.
Qy	180	ADEGM--DCALAAVYGGKGVLSQLRELKAWFEEQTLFHYFASILLGYDAAV-AAGG 236	PR	01-JUN-1999;	99US-0137222P.
Db	179	ADSNLTNCAPASEVYGGCGILAQLELKDWFETQTLYHFNCSILMIYENESILMQGG 238	PR	03-JUN-1999;	99US-0137528P.
Qy	237	DGG---GVTVKLVDFAHVAGDGVLDHNFGLGSLIKFVSDIV 277	PR	04-JUN-1999;	99US-0137502P.
Db	239	DDAPAPRAQKLVDFAHVLDGNGVIDHNFGLGSLFIKFIKDIL 282	PR	07-JUN-1999;	99US-0137724P.
RESULT 8					
ID	AAG46590 standard; protein; 330 AA.				
XX	AC	AAG46590;			
XX	DT	18-OCT-2000 (first entry)			
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 58630.			
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;			
XX	KW	hybridisation assay; genetic mapping; gene expression control; promoter;			
XX	KW	termination sequence.			
XX	OS	Arabidopsis thaliana.			
XX	PN	EP1033405-A2.			
XX	PD	06-SEP-2000.			
XX	PF	25-FEB-2000; 2000EP-00301439.			
XX	XX				
PR	25-FEB-1999;	99US-0121825P.	PR	13-JUL-1999;	99US-0143542P.
PR	05-MAR-1999;	99US-0123180P.	PR	14-JUL-1999;	99US-0143624P.
PR	09-MAR-1999;	99US-0123548P.	PR	15-JUL-1999;	99US-0144005P.
PR	23-MAR-1999;	99US-0125788P.	PR	16-JUL-1999;	99US-0144085P.
PR	25-MAR-1999;	99US-0126264P.	PR	18-JUL-1999;	99US-0144086P.
PR	29-MAR-1999;	99US-0126785P.	PR	19-JUL-1999;	99US-0144325P.
PR	01-APR-1999;	99US-0127462P.	PR	19-JUL-1999;	99US-0144331P.
PR	06-APR-1999;	99US-0128234P.	PR	19-JUL-1999;	99US-0144332P.
PR	08-APR-1999;	99US-0128714P.	PR	19-JUL-1999;	99US-0144333P.
PR	16-APR-1999;	99US-0129845P.	PR	19-JUL-1999;	99US-0144334P.
PR	19-APR-1999;	99US-0130077P.	PR	20-JUL-1999;	99US-0144352P.
PR	21-APR-1999;	99US-0130449P.	PR	20-JUL-1999;	99US-0144632P.
PR	23-APR-1999;	99US-0130510P.	PR	20-JUL-1999;	99US-0144884P.
PR	23-APR-1999;	99US-0130891P.	PR	21-JUL-1999;	99US-0144814P.
PR	28-APR-1999;	99US-0131449P.			

PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161359P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146388P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161920P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161992P.
PR 04-AUG-1999;	99US-0147204P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147302P.	PR 29-OCT-1999;	99US-0162142P.
PR 05-AUG-1999;	99US-0147312P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 08-AUG-1999;	99US-0147493P.		
PR 08-AUG-1999;	99US-0147935P.		
PR 11-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 15-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 20-AUG-1999;	99US-0149929P.		
PR 23-AUG-1999;	99US-0149802P.		
PR 23-AUG-1999;	99US-0149930P.		
PR 23-AUG-1999;	99US-0150566P.		
PR 26-AUG-1999;	99US-0150884P.		
PR 27-AUG-1999;	99US-0151065P.		
PR 27-AUG-1999;	99US-0151066P.		
PR 27-AUG-1999;	99US-0151080P.		
PR 30-AUG-1999;	99US-0151303P.		
PR 31-AUG-1999;	99US-0151438P.		
PR 01-SEP-1999;	99US-0151930P.		
PR 07-SEP-1999;	99US-0152363P.		
PR 10-SEP-1999;	99US-0153070P.		
PR 13-SEP-1999;	99US-0153758P.		
PR 15-SEP-1999;	99US-0154018P.		
PR 16-SEP-1999;	99US-0154039P.		
PR 20-SEP-1999;	99US-0154779P.		
PR 22-SEP-1999;	99US-0155139P.		
PR 23-SEP-1999;	99US-0155486P.		
PR 24-SEP-1999;	99US-0155659P.		
PR 28-SEP-1999;	99US-0156458P.		
PR 29-SEP-1999;	99US-0156596P.		
PR 04-OCT-1999;	99US-0157117P.		
PR 05-OCT-1999;	99US-0157753P.		
PR 06-OCT-1999;	99US-0157855P.		
PR 07-OCT-1999;	99US-0158029P.		
PR 08-OCT-1999;	99US-0158232P.		
PR 12-OCT-1999;	99US-0158369P.		
PR 13-OCT-1999;	99US-0159293P.		
PR 13-OCT-1999;	99US-0159294P.		
PR 13-OCT-1999;	99US-0159295P.		
PR 14-OCT-1999;	99US-0159329P.		
PR 14-OCT-1999;	99US-0159330P.		
PR 14-OCT-1999;	99US-0159331P.		
PR 14-OCT-1999;	99US-0159637P.		
PR 14-OCT-1999;	99US-0159638P.		
PR 18-OCT-1999;	99US-0159884P.		

Query Match 47.0%; Score 714.5; DB 3; Length 330;

Best Local Similarity 52.8%; Pred. No. 8.4e-70;

Matches 150; Conservative 39; Mismatches 82; Indels 13; Gaps 8;

QY 4	LHPPHQVAGHRASAKLGLIDSGLFYKPLQAGDRGEHEVAFYAFSAHAAVAPARID	63
DB 32	LKVPEHQVAGHTIASDGKLGPLVDDQGRFFKPLQAGDRGEHEAFYEFSTSNMKVPDHIH-	90
QY 64	TFEPREHGTLLPTEAQPGEHPH-PLVLDDLLAGFOACVADIKIGAITWPPSSPEYIA	122
DB 91	RYFVYHGTQLV--EASDGGKPLHLVDDVDSGYANPSVMDKIGRTWPDVSEYFK	148
QY 123	KCLAKDRGTTSVLLGFRVSGRVVG-PEGAVMTERPEVKAMDTAGVRRVLRYS--SV	179
DB 149	KCIKDRQTTVSLGFRVSGFKIFDHQESSFWAEKKLVLYGNADGALALRKFVSSNSP	208
QY 180	ADEGM--DCALAAVYGGKGVLSQURELKANPEOTLPHFYASILLGVDAAV-AAG	236
DB 209	ADSNLTFCNAPASEVYGGCNGILAQLELKDWPETQTLXHFNSCSILMIYENESILMQG	268
QY 237	DGG--GVTVKLVDFAHVAEGDGVIDHNFLLGGICSLIKFVSDIV	277
DB 269	DDAPAPRAQVKGVDFARVLDGNGVIDHNFLLGGICSLIKFVSDIV	312

RESULT 9

AAG24257

ID AAG24257 standard; protein; 300 AA.

XX AAG24257;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 27860.

XX Arabidopsis thaliana

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-01334256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135829P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 18-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140895P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.

```
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159338P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161820P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.1%; Score 700.5; DB 3; Length 300;
Best Local Similarity 51.8%; Pred. No. 2.6e-68;
Matches 147; Conservative 41; Mismatches 83; Indels 13; Gaps 8;

QY 4 LHPPEHQVAGHRASAKGLPLIDSGLFYKPLQAGDRGHEVAFYAFSAHAAPARID 63
Db 2 LKVPHQVAGHTASDKGLPLVDQGRFPKPLQDSRGHEAKFYESTSNKKVPDHIH- 60

QY 64 TFPFPHGTRLLPTEAQGPETH-PHVLDDLLAGFOAPCVADIKIGAITWPPSSPEFYA 122
Db 61 RYFPVYHGTLV--EASDGSGLPHLVDVVSGYANPSVMDVKIGSRWTYPDVSEYFK 118

QY 123 KCLAKDRGTSVLLGFRVSGVRVVG-PEGAVRTERPEVKAMDTAGVRRLRYVS--SV 179
Db 119 KCIKEDRQXTTVSLGFRVSGFKIFDQESFWRAEKKLVLYGNADGARLAKRFVSSNP 178

QY 180 ADEGM--DCALAAVYGGKGVLSQLRELKAWFEOTLFHYFSASILLGYDAAV-AAGG 236
Db 179 ADSNLTFCAFASEVYGGCGNGLAQLELKDWFETQTLHFNSCSILMIYENESILMKGG 238

QY 237 DGG--GVTVKLVDPFAHVAEGDVIDHNFGLGLCSLIKVFSDIV 277
Db 239 DDAPAPRAQVKLVNFAHVLDCNGVIDHNFGLGLCSFIKIKDIL 282

RESULT 10
AAG24256
ID AAG24256 standard; protein; 330 AA.
XX
AC AAG24256;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27859.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.

XX
PR 25-FEB-2000; 2000EP-00301439.
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142053P.
PR 06-JUL-1999; 99US-0142390P.
```









```
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0142830P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144311P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 27-JUL-1999; 99US-0145319P.
PR 28-JUL-1999; 99US-0145319P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149368P.
PR 18-AUG-1999; 99US-0149175P.
PR 20-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.

PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149910P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 23-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 44.3%; Score 674.5; DB 3; Length 286;
Best Local Similarity 50.4%; Pred. No. 1.8e-65;
Matches 141; Conservative 42; Mismatches 86; Indels 11; Gaps 7;

QY 4 LHPPEHQVAGHRASAKLGPLIDGSLFYKPLQAGDRGHEVAFYEAFAHAAVAPARID 63
Db 3 LKVPHEQVAGHIAKDGKPCPLVDDKGRFKPLQGSRGELFKVFEFSNTVEPHIH- 61
QY 64 TFFRPHGTRLLPTEAQPGCEPHPHVLVDLLAGFQAPCVADIKIGAITWPPSPPEYIAK 123
Db 62 RYFVYHGTQAV--EGSDGA--AMVLENLLAEYSKPSVMDVMGSRWTWYDASEYIQK 117
QY 124 CLAKDRGTTSVLLGFRVSGVRVVG--PEGAVWTERPEVXAMDTAGVRVLRVYVS--SVA 180
Db 118 CLKODTGTITVSSGRFISGFVEYDHKESFVKPKRLRGLDVGARLTLRKFEVSSNLS 177
QY 181 DEGM--DCALAAAVYGGKGVLSQLRELKAMFEBQTLPHFYASILLGYDAAAAGGDG 238
```

```
Db      178 DTGSKPDSAFSSVYGGSHGLTQLLEKWTWFENQTLHNSCSILMVYENESILKGNDD 237
Qy      239 GG-VTVKLVDFAHVAEGDVIDHNFGLGSLIKFVSDIV 277
Db      238 DARPQVKLVDFAHVLDGNGVIDHNFGLGSLCFINFIREFIL 277

RESULT 13
AAG06976
ID      AAG06976 standard; protein; 286 AA.
XX
AC      AAG06976,
XX
DT      17-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana protein fragment SEQ ID NO: 3951.
XX
KW      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
XX      termination sequence.
XX
OS      Arabidopsis thaliana.
XX
FN      EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PF      25-FEB-2000; 2000EP-00301439.
XX
PR      25-FEB-1999; 99US-0121825P.
PR      05-MAR-1999; 99US-0123180P.
PR      09-MAR-1999; 99US-0123548P.
PR      23-MAR-1999; 99US-0125788P.
PR      25-MAR-1999; 99US-0126284P.
PR      29-MAR-1999; 99US-0126785P.
PR      01-APR-1999; 99US-0127462P.
PR      06-APR-1999; 99US-0128234P.
PR      08-APR-1999; 99US-0128714P.
PR      16-APR-1999; 99US-0129845P.
PR      19-APR-1999; 99US-0130077P.
PR      21-APR-1999; 99US-0130449P.
PR      23-APR-1999; 99US-0130510P.
PR      23-APR-1999; 99US-0130891P.
PR      28-APR-1999; 99US-0131449P.
PR      30-APR-1999; 99US-0132048P.
PR      30-APR-1999; 99US-0132407P.
PR      04-MAY-1999; 99US-0132484P.
PR      05-MAY-1999; 99US-0132485P.
PR      06-MAY-1999; 99US-0132486P.
PR      07-MAY-1999; 99US-0132487P.
PR      07-MAY-1999; 99US-0132863P.
PR      11-MAY-1999; 99US-0134256P.
PR      14-MAY-1999; 99US-0134218P.
PR      14-MAY-1999; 99US-0134219P.
PR      14-MAY-1999; 99US-0134221P.
PR      18-MAY-1999; 99US-0134370P.
PR      18-MAY-1999; 99US-0134768P.
PR      19-MAY-1999; 99US-0134941P.
PR      20-MAY-1999; 99US-0135124P.
PR      21-MAY-1999; 99US-0135353P.
PR      24-MAY-1999; 99US-0135629P.
PR      25-MAY-1999; 99US-0136021P.
PR      27-MAY-1999; 99US-0136392P.
PR      28-MAY-1999; 99US-0136782P.
PR      01-JUN-1999; 99US-0137222P.
PR      03-JUN-1999; 99US-0137528P.
PR      04-JUN-1999; 99US-0137502P.
PR      07-JUN-1999; 99US-0137724P.
PR      08-JUN-1999; 99US-0138094P.
PR      10-JUN-1999; 99US-0138540P.
PR      10-JUN-1999; 99US-0138847P.
PR      14-JUN-1999; 99US-0139119P.
PR      16-JUN-1999; 99US-0139452P.
PR      16-JUN-1999; 99US-0139453P.
PR      17-JUN-1999; 99US-0139452P.
PR      18-JUN-1999; 99US-0139454P.
PR      18-JUN-1999; 99US-0139455P.
PR      18-JUN-1999; 99US-0139456P.
PR      18-JUN-1999; 99US-0139457P.
PR      18-JUN-1999; 99US-0139458P.
PR      18-JUN-1999; 99US-0139459P.
PR      18-JUN-1999; 99US-0139460P.
PR      18-JUN-1999; 99US-0139461P.
PR      18-JUN-1999; 99US-0139462P.
PR      18-JUN-1999; 99US-0139463P.
PR      18-JUN-1999; 99US-0139750P.
PR      18-JUN-1999; 99US-0139763P.
PR      21-JUN-1999; 99US-0139817P.
PR      22-JUN-1999; 99US-0139899P.
PR      23-JUN-1999; 99US-0140353P.
PR      23-JUN-1999; 99US-0140354P.
PR      24-JUN-1999; 99US-0140695P.
PR      28-JUN-1999; 99US-0140823P.
PR      29-JUN-1999; 99US-0140991P.
PR      30-JUN-1999; 99US-0141287P.
PR      01-JUL-1999; 99US-0141842P.
PR      01-JUL-1999; 99US-0142154P.
PR      02-JUL-1999; 99US-0142055P.
PR      06-JUL-1999; 99US-0142390P.
PR      08-JUL-1999; 99US-0142803P.
PR      09-JUL-1999; 99US-0142920P.
PR      12-JUL-1999; 99US-0142977P.
PR      13-JUL-1999; 99US-0143542P.
PR      14-JUL-1999; 99US-0143624P.
PR      15-JUL-1999; 99US-0144005P.
PR      16-JUL-1999; 99US-0144085P.
PR      16-JUL-1999; 99US-0144086P.
PR      17-JUL-1999; 99US-0144325P.
PR      19-JUL-1999; 99US-0144331P.
PR      19-JUL-1999; 99US-0144332P.
PR      19-JUL-1999; 99US-0144333P.
PR      19-JUL-1999; 99US-0144334P.
PR      19-JUL-1999; 99US-0144335P.
PR      20-JUL-1999; 99US-0144352P.
PR      20-JUL-1999; 99US-0144632P.
PR      20-JUL-1999; 99US-0144884P.
PR      21-JUL-1999; 99US-0144814P.
PR      21-JUL-1999; 99US-0145086P.
PR      21-JUL-1999; 99US-0145088P.
PR      22-JUL-1999; 99US-0145085P.
PR      22-JUL-1999; 99US-0145087P.
PR      22-JUL-1999; 99US-0145089P.
PR      22-JUL-1999; 99US-0145192P.
PR      23-JUL-1999; 99US-0145145P.
PR      23-JUL-1999; 99US-0145218P.
PR      23-JUL-1999; 99US-0145224P.
PR      26-JUL-1999; 99US-0145276P.
PR      27-JUL-1999; 99US-0145913P.
PR      27-JUL-1999; 99US-0145918P.
PR      27-JUL-1999; 99US-0145919P.
PR      28-JUL-1999; 99US-0145951P.
PR      02-AUG-1999; 99US-0146386P.
PR      02-AUG-1999; 99US-0146388P.
PR      02-AUG-1999; 99US-0146389P.
PR      03-AUG-1999; 99US-0147038P.
PR      04-AUG-1999; 99US-0147204P.
PR      04-AUG-1999; 99US-0147302P.
PR      05-AUG-1999; 99US-0147192P.
PR      05-AUG-1999; 99US-0147260P.
PR      06-AUG-1999; 99US-0147303P.
PR      06-AUG-1999; 99US-0147416P.
PR      09-AUG-1999; 99US-0147493P.
PR      09-AUG-1999; 99US-0147935P.
PR      10-AUG-1999; 99US-0148171P.
PR      11-AUG-1999; 99US-0148319P.
```



PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142772P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.

PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.



Query Match 44.3%; Score 674.5; DB 3; Length 327;  
Best Local Similarity 50.4%; Pred. No. 2.2e-65;  
Matches 141; Conservative 42; Mismatches 86; Indels 11; Gaps 7;  
QY 4 LHPPEHQVAGHRASAKLGPLIDSGGLFYKPLQAGDGEHEVAFYEAFAHAAPARIRD 63  
Db 44 LKYPEHQVAGHIAKDGKPGPLVDKGRFFKPLQDGRGEIEVKFYBSFSSNTEVPEHIH- 102  
QY 64 TFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSPPEPIAK 123  
Db 103 RYFPVYHGTQAV--EGSDGA--AMVLENLLAAYSKSVMDVKMGSTWYTPDASEEYIQ 158  
QY 124 CLAKDRGTTSVLLGFRVSGVRVVG-PEGAVWRTERPEVKAMDAGVRRVLRVYS--SVA 180  
Db 159 CLKKDTGTTTSSGFRISGFVVDHKESFVKPERKLLRGLDVGDGARLTLRKFVSSNLS 218  
QY 181 DEGM--DCALAAAVYGGKGVLSQRELKAWFEQTLFHFYSASILLGYDAAAAGG 238  
Db 219 DTGSKPDSAFASVYGGSHGILITQLLELKTWFENQTLHFNSCSILMVYENESILK 278  
QY 239 GG-VTVKLVDFAHVARGDGVIDHNFGLGSLIKFVSDIV 277  
Db 279 DARPQVKLVDFAHVLDGNGVIDHNFGLGSLFINFIREIL 318

RESULT 15  
ABB93390  
ID ABB93390 standard; protein; 286 AA.  
AC ABB93390;  
XX  
DT 31-MAY-2002 (first entry)  
XX  
DE Herbicidally active polypeptide SEQ ID NO 2601.  
XX  
KW Herbicidal; plant; agriculture; herbicide.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200210210-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 28-AUG-2001; 2001WO-EP009892.  
XX  
PR 28-AUG-2001; 2001WO-EP009892.  
XX  
PA (FARB ) BAYER AG.  
XX  
PI Tietjen K, Weidler M;  
XX  
PT WPI; 2002-269010/31.  
XX  
DR Identifying plant target proteins for herbicidally active compounds,  
XX comprising aligning and comparing nucleic acid or amino acid sequences  
XX from plant with nucleic acid or amino acid sequences from non-plant  
XX organisms.  
XX  
PS Claim 5; SEQ ID NO 2601; 261pp + Sequence Listing; English.  
XX  
CC The invention relates to identifying target proteins (ABB90790-ABB94016)  
XX for herbicidally active compounds, comprising aligning and comparing  
XX nucleic acid or amino acid sequences from plant with nucleic acid or  
XX amino acid sequences from non-plant organisms using suitable search  
XX parameters, where plant sequences having an E-value greater by a factor  
XX of 3 than the E-value of most similar non-plant sequences are selected.  
XX the polypeptides or nucleic acids encoding them are useful for  
XX identifying modulators. The identified modulators are useful as  
XX herbicides  
XX  
SQ Sequence 286 AA;

Query Match 44.3%; Score 673.5; DB 5; Length 286;  
Best Local Similarity 50.4%; Pred. No. 2.3e-65;  
Matches 141; Conservative 42; Mismatches 86; Indels 11; Gaps 7;  
QY 4 LHPPEHQVAGHRASAKLGPLIDSGGLFYKPLQAGDGEHEVAFYEAFAHAAPARIRD 63  
Db 3 LKYPEHQVAGHIAKDGKPGPLVDKGRFFKPLQDGRGEIEVKFYBSFSSNTEVPEHIH- 61  
QY 64 TFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSPPEPIAK 123  
Db 62 RYFPVYHGTQAV--EGSDGA--AMVLENLLAAYTKPSVMDVKMGSTWYTPDASEEYIQ 117  
QY 124 CLAKDRGTTSVLLGFRVSGVRVVG-PEGAVWRTERPEVKAMDAGVRRVLRVYS--SVA 180  
Db 118 CLKKDTGTTTSSGFRISGFVVDHKESFVKPERKLLRGLDVGDGARLTLRKFVSSNLS 177  
QY 181 DEGM--DCALAAAVYGGKGVLSQRELKAWFEQTLFHFYSASILLGYDAAAAGG 238  
Db 178 DTGSKPDSAFASVYGGSHGILITQLLELKTWFENQTLHFNSCSILMVYENESILK 237  
QY 239 GG-VTVKLVDFAHVARGDGVIDHNFGLGSLIKFVSDIV 277  
Db 238 DARPQVKLVDFAHVLDGNGVIDHNFGLGSLFINFIREIL 277

Search completed: March 24, 2004, 20:57:40  
Job time : 63 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 20:56:29 ; Search time 23 seconds  
(without alignments)  
648.692 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPRHQVAGHRASK.....IKFVSIVPETHPTQLGPS 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	96.5	6.3	502	4	US-09-489-039A-14063
2	96	6.3	631	4	US-09-252-991A-27902
3	94.5	6.2	408	4	US-09-252-991A-32393
4	94	6.2	418	4	US-09-489-039A-14152
5	93.5	6.1	384	4	US-09-252-991A-31358
6	92.5	6.1	882	4	US-09-252-991A-17653
7	92	6.0	795	4	US-09-252-991A-19085
8	91.5	6.0	424	4	US-09-252-991A-21828
9	91.5	6.0	1175	4	US-09-252-991A-25044
10	90.5	6.0	926	4	US-09-252-991A-31053
11	90	5.9	311	4	US-09-252-991A-29142
12	90	5.9	677	4	US-09-252-991A-20406
13	90	5.9	1567	4	US-09-252-991A-20477
14	89	5.9	409	4	US-09-252-991A-31351
15	89	5.9	4551	3	US-09-320-878-1
16	89	5.9	4551	4	US-09-141-908-2
17	89	5.9	4551	4	US-09-657-440-1
18	89	5.9	4613	3	US-09-105-537-31
19	89	5.9	11877	3	US-09-105-537-6
20	88.5	5.8	370	2	US-08-846-762-81
21	88.5	5.8	474	4	US-09-489-039A-8680
22	88.5	5.8	710	4	US-09-287-849-16
23	88.5	5.8	856	4	US-09-287-849-12
24	88	5.8	463	4	US-09-252-991A-29731
25	88	5.8	485	4	US-09-252-991A-31436
26	87.5	5.8	407	2	US-08-926-327-2
27	87.5	5.8	407	3	US-09-119-918-2

#### ALIGNMENTS

##### RESULT 1

US-09-489-039A-14063  
; Sequence 14063, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 14063  
; LENGTH: 502  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-14063

Query Match Similarity 6.3%; Score 96.5; DB 4; Length 502;  
Best Local Similarity 21.4%; Pred. No. 0.061; 98; Indels 117; Gaps 14;  
Matches 68; Conservative 35; Mismatches 35

QY	42	EHEVAFYBAFSA--HAAVPAIRIDTFFRFGHTRLLPTAQGPPEPHLVLDLLAGFQA	99
DB	28	ENTVVFPEAFIAQMEAMP	60
QY	100	PCVADIKIGAT-----WPPSPPEPIAKLAKDR-----GTTSVLLG	137
DB	61	PIRRSIRVTLTKISVDDELFLVAPYGV--QLAPVPWCAGFWIEREDDDALPLGSTAEHLS	119
QY	138	--FRVSGVVVPEGAVWRTERPEVKAMDAGVRRLRVYSSVADEGDCALAAVYGG	195
DB	120	GLIFYQESAMPVVAALFADNRQPERVMDVA-----AAGSKTTQIAARMGN	166
QY	196	KGGVL-----SQRLEKA-----WFERQTLFHF-----YSASILGYDAAAAGGDDGG	240
DB	167	AGGILANFSASRVKVLHNSRCGISNVALTDFDGRVFGAALPETFDAILLDAPCSGEG	226
QY	241	VTVK-----LVDFV--HVAAGGVGDHNFGLGLC-----SL	269
DB	227	VVRKQADALKNWPDSNLDIAATQRELIDSFAHALRPGGTLVYS-----TCTLNREENOSV	282
QY	270	IKFVSIVPETHPTQLG	287
DB	283	IEWLLSRYPQAVEILPLG	300

##### RESULT 2

```
US-09-252-991A-27902
; Sequence 27902, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27902
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27902

Query Match 6.3%; Score 96; DB 4; Length 631;
Best Local Similarity 23.6%; Pred. No. 0.098;
Matches 70; Conservative 30; Mismatches 82; Indels 114; Gaps 16;

QY 76 PTEAQGPBPHLVLDLLAGFOAPCVADIKIGAITWPPSSPEPIYAKC---LAKD---R 129
Db 333 PDAAGADPGEQLV-----RCLAFVE---TDPVAPFAFGTCLRLAEQVPFV 378
QY 130 GTTSVL-----LGRVSGRVVVGEGAVWTERPEVKAMDTAGVRR---VLRVYSS 178
Db 379 GDSALVEPRGRFVLAHRFQAHGAV-----TVDPEVDGYAIHQRRHQIVGQGLAS 432
QY 179 VA-----DEGMDCALAAAVYGGKGVV-----LSQLREL 206
Db 433 VAAAFQDQRRRQRRQRQAVDQGVQVLDLAI-APVHVGVGGLRRALAGVEAIGRLRG 491
QY 207 KAWFEQTLFHFVSASILGY-----DAAAVAAAG-----GDGG--- 239
Db 492 LAQKEESTAAE--SRAVLRYGQSGDGRRSVEGVAAGQGFQAGDGRGMRMGDGLLR 549
QY 240 --GVTVKLVDFAHVAE-----GDGVLDHNFVGGCLSLKFKVSDIVPEPHT 293
Db 550 CVGVAAKAVGMRREARQGAAGDQGTGQVQBGHGSLSFGVGRGTG---PRTPQT 602

RESULT 3
US-09-252-991A-32393
; Sequence 32393, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32393
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32393

Query Match 6.2%; Score 94.5; DB 4; Length 408;
Best Local Similarity 23.6%; Pred. No. 0.073;
Matches 57; Conservative 27; Mismatches 93; Indels 65; Gaps 10;

QY 53 AHAAPARIRDTFFPHGTRLLPTTEAQGP-----HPLVLDDLAGE----- 97

US-09-252-991A-27902
; Sequence 27902, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27902
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27902

Query Match 6.2%; Score 94; DB 4; Length 418;
Best Local Similarity 25.2%; Pred. No. 0.087;
Matches 54; Conservative 20; Mismatches 70; Indels 70; Gaps 11;

QY 32 YKPLQAGDGRGEHEVAFYFAFSAHAAVPAKIRDTFFPHGTRLLPTTEAQGPBPHLVLD 91
Db 230 YQPLSA---EEIFAFYEVCVRHVSVPVCLYDN--PR-----TTHVMLA 267
QY 92 DLLAGFOA--PCVADIKIGAITWPPSSPEPIYAKLAKDRGTTSVLLGFRVSGVR----- 144
Db 268 DELQGRIRALPAIASIKI-----FCLPAPQASE-----RVAALRQHLP 306
QY 145 --VVGPEGAVWTERPEVKAMDTAGVRRVLRVYSSVADEGMDCALA--AAVYGGKGVV 200
Db 307 RVTLGVSGDAWA-----TAGLQAGCEAWYVSGGLFPHCSLALVRAIRSGDVAQT 356
QY 201 SQLRELKA--WFEEQTLFHFYSASILLGYDAAA 232
Db 357 AALNEQLAPLW---RCFDRYGGSLRVITASAAA 386

RESULT 5
US-09-252-991A-31358
; Sequence 31358, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
```

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31358
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31358

Query Match
Best Local Similarity 6.1%; Score 93.5; DB 4; Length 384;
Matches 73; Conservative 37; Mismatches 124; Indels 107; Gaps 15;

QY 5 HPPEHQVAG-HRASASKLGLIDG-----SGLFYKPIQ----- 36
Db 82 HPGQRLGFGHALGCGPADLVHQVLREVLALEAARAEITTVVGGEVFTLQAGKETS 141
QY 37 ----AGDRGEHEVAFYAFSAHAAPAKIRDTFFPRFHGTRLLPTEAQPGEPPHVLDD 92
Db 142 AQRVGDGDAQAT---AGGHPGV-----LRIAGPQVRLALQADRVRHPVGTLEG 189
QY 93 LLAGFQAPCVADI----KIGAITWPPSPPEPIAKCLAKDRGTTSVLLGF-RVSGVRVVG 147
Db 190 FRGFAETDVADLAGLHQVG-----EGADAVLDGLYLRVHAVLIVQ 229
QY 148 PEGAVWRTEREVKAMDTPAGVRR-----VLRVYVSSVADE-----GMDCLAAAVYGGKGV 199
Db 230 VQAV--GTEALQAPDGAADVRAAVDAARPGIGRVADDAELAGQEHLLAFAAQGLAEQF 287
QY 200 LSQRELKAWFEEQTLFHFYSA-----SILLGYDAAVAAGDGGGVTVKLV 246
Db 288 LVTVRVHVGGIEEVQPEFHGMQGGDGLLVAAGGVGVGHAHAARANGDGRVAVAEIT 347
QY 247 DFAHVAEGDGVIDHNFLGG-----LCSLIKFSVDIVETP 281
Db 348 GF-HVS-----SLAGGKENVILRSALVWIKPIITVLP 379

RESULT 6
US-09-252-991A-17653
; Sequence 17653, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17653
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (730)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-17653

Query Match
Best Local Similarity 6.1%; Score 92.5; DB 4; Length 882;
Matches 85; Conservative 29; Mismatches 99; Indels 109; Gaps 22;

QY 11 VAGHRASAKLGPL-IDSGLYFKPLQAGDRGEHEVAFYAFSAHAAPAKIRDTFFPRF 69
Db 128 VAGHQAEARVAGDLAVDAB-----QEAADR-----NPVAFPAVAAL-----GTLAP-- 168

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31358
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31358

Query Match
Best Local Similarity 21.4%; Pred. No. 0.086;
Matches 73; Conservative 37; Mismatches 124; Indels 107; Gaps 15;

QY 70 HGTRLLP-----TEAQPGE---PPHVLVLDLLE- 97
Db 169 GAERLEFAFGSLPAGLAVAGLAQAFADVAQIAAVVEQAGERRVDFGQLLEGFLGCG 228
QY 98 ---QAPCVADIKIGAITWPPSPPEPIYAKCLAKDRGT-TSVLLGFRVSGV-RVVGPEGAV 152
Db 229 VFRQEAFVDRVAGAEFRPHGTSRRLLVGR--LRDAGEGRAVASDIQADGVGRVV--QGAA 284
QY 153 -----WRTERPEVKAMDTPA-GVRRVLRVRYVSSVADEGMDCLAA---AAYVGGKGVLS 201
Db 285 EEQVVRQVRYVEAP---LDAAGGVEDVV-----AAADQALDLGLAEQVRAVFRHGGQDIAA 335
QY 202 Q-----LREL-----KAWFEEQTLFHFYSAIILGYDAAVAAGDGGGVTV 243
Db 336 RDARLAGDDPVGAFKRLHRVGRGFEAQV-----PGDLOGAD--GVARGHRAAGARQ 387
QY 244 KLVDFAHVAEGDGVIDHNFLGG 265
Db 388 QAADPAVAEGAAGID--FDGG 407

RESULT 7
US-09-252-991A-19085
; Sequence 19085, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19085
; LENGTH: 795
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19085

Query Match
Best Local Similarity 27.3%; Pred. No. 0.4; Length 795;
Matches 51; Conservative 16; Mismatches 70; Indels 50; Gaps 12;

QY 6 PPEHQV-----AGHRASAKLGLIDGSLFYKPLQAGDRGEHEVAFY---EAFSAHA 55
Db 188 PPVHRFRHGPVRSFGRRAGQG-LGILLGADGHHHR-LRRGQGHRTVHHPTDEAFRLQA 245
QY 56 AVPARIRDTFFPRFHGTRLLPTEA-----QFGEPPHVLVLDLLAGFOAPCVADIKI 107
Db 246 GVRERRRG---DREHGQ---PTDAAGDGGGVHYGRDQRELRGDRQGGAD-PGSAVFR 298
QY 108 GAITWBPSSPEPIYAKCLAKDRG---TTSVLLGFR-----VSG---VRVVGPEGA 151
Db 299 GVLDCGPRGPP-----GRSKGPAARSVERLRCRPRALVPADSVAGADRAAULRAHA 351
QY 152 VWRTERP 158
Db 352 VLRNHRP 358

RESULT 8
US-09-252-991A-21828
; Sequence 21828, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
```



; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 29142  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29142

Query Match 5.9%; Score 90; DB 4; Length 311;  
Best Local Similarity 25.6%; Pred. No. 0.15;  
Matches 54; Conservative 25; Mismatches 72; Indels 60; Gaps 12;  
QY 7 PEHQVAGHRASAKGLPLIDGSLFYKPLQAGDRGEHEVAFYAFSAHAAPVPA-----RI 61  
Db 84 PAPPAPCRAGGLALAGPAVPRG-----AGRAGLAEVAPTQSLARPAPVAGGRQPRL 136  
QY 62 RDTFFP--RPHGTRLLPTEAQGP-----HPHL-VLDDL----- 93  
Db 137 RPAYRPAGRQRRPMDRQRPGLFPQFLPQSRSLRQSPYGLHVAVDVRRGARRVAR 196  
QY 94 LAGFQAPCVADIKIGAITWPPSPPEPIAKLAKDRGTTSVL--LGRFVSQVVRVGP 151  
Db 197 LADHPPPPVADARHDAARRRAEP-----DRSLAALGAAGHRLAGTEAIA--GA 244  
QY 152 VWRTERPEVKAMDTAGVRRVL--RRYVSSVA 180  
Db 245 VPR--RAVLPLLEA--RRLIEPRRHGCV 271

## RESULT 12

US-09-252-991A-20406  
; Sequence 20406, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20406  
; LENGTH: 677  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20406

Query Match 5.9%; Score 90; DB 4; Length 677;  
Best Local Similarity 28.7%; Pred. No. 0.52;  
Matches 64; Conservative 18; Mismatches 69; Indels 72; Gaps 15;  
QY 5 HPPEHVOAGHRASAKGLPLIDGSLFYKPLQAGDRGEHEVAFYAFSAHAAPVPAIRD 63  
Db 31 HPARHLLFLLRAGCAIAAPVQGRGAW-----RGQRPVDPAGVRRAGAGH-----RHRR 80  
QY 64 TFPFRPHGTRLLPTEAQGPPEPHPLVLDDLQAGFAP-----CVADIKIGAIT--W 112  
Db 81 GGVQRQH--RDLF-EAVPGAP-----ERLAGVAADPRGGQGLC--LRGSAIDHG 127  
QY 113 PPS-SPPEPIAKLAKDRGTTSVLLGFRVSGVRVGP-----EGAVWRTERPEVKAMD 164  
Db 128 PPRSPEPEVA-----VRLPHLLFGAAAGFLRPRQRSGAVR-----AVH 169  
QY 165 TAGVR-----RVLRVYSSVADGMDCALAAAVYGGKG 198  
Db 170 VRGVPRGDCRCDRPLRR--AAVAGNGLRPAFLFPVVRROGG 210

## RESULT 13

Query Match 5.9%; Score 89; DB 4; Length 409;

US-09-252-991A-20477  
; Sequence 20477, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20477  
; LENGTH: 1567  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20477

Query Match 5.9%; Score 90; DB 4; Length 1567;  
Best Local Similarity 24.5%; Pred. No. 1.9;  
Matches 80; Conservative 37; Mismatches 114; Indels 96; Gaps 19;  
QY 9 HQVAGHRASAKGLPLIDGSLFYKPLQAGDRGE-----LFFKPLQAGDRGE-----HEVAFYEA 50  
Db 1075 HEITGG---GAGVLGG--DGDGRAFOQVRVLADRLQLAQGEAQEQDQADHQHQRHRADEQV 1129  
QY 51 FSAHAAVPAIRD-----TFPRPHGTRLLPTEAQGPPEPHPLVLDDLQAGFQA-- 99  
Db 1130 GKHRGVPSPAAGDRHGRGAAGVFLDRFHGHHPVAQVDSL--GSHYP--VALGEALDFHAT 1187  
QY 100 -PCVADIKIGAITWPPSPPEPIAKLAKDRGTTSVLL--GPRVSGVRVWG-----PEGA 151  
Db 1188 VPALADPHRDAL-----GAERPVAVLVGDHGVAVGGIGDGRGFRHQAQRRVGFARDPD-- 1240  
QY 152 VWRTERPEVKAMDTAG-----VRRVLRVYSSVADGMDCALAAAVYGGKGCV--LS 201  
Db 1241 -----AGEHAGGEAQAFVGVQGLDHLHRLADQVRVDGADARREAPRIGVDQDLE 1290  
QY 202 QLRELK---AWFEQTLF-----HFYSASILLGVDAVAAGDGGGVTVKLVLD 247  
Db 1291 GLAELQLGTGAFVQGEIAFCGPRIHQAGHRLGDVAEVLAEVHAADAEELAGEGRADRL--LVD 1349  
QY 248 -----FAHVAEGDGVIDHNLFGG 265  
Db 1350 QRLGLAHRAEHRVAGDGVV--QLFLRG 1375

## RESULT 14

US-09-252-991A-31351  
; Sequence 31351, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 31351  
; LENGTH: 409  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-31351

Best Local Similarity 22.9%; Pred. No. 0.3;  
Matches 83; Conservative 24; Mismatches 90; Indels 166; Gaps 19;  
Qy 6 PPEHQVAGHR-----ASASKLGLIDSGSLFYKPLQAGD-----RGEH 43  
Db 33 PPRQPARSRAVPDRSLDGOAVRLG---DGRG-----SLLRGDDPRWLMLPQGLPGGTH 85  
Qy 44 EVAEYEAFAHAAPAIRDTFFPRFHGTRLLPTE-AOP-----81  
Db 86 -----LSGRAGLPDL-----RGHRAAPGDRAPVAATIRPFNGAFSLQGEAL 131  
Qy 82 ---GEPHPLV-----LDDLLAGFQAPCVADIKIGAITWP 113  
Db 132 RHLGSKCPSLVRPTPLGRSRRLRLVARGDLALQDLILA---HFAVGRVLIGAL--- 185  
Qy 114 PSSPEPIAKLAKDRGTTSVLLGFRVSGV-----RVVG---PEGAVRTERPEVKAMDTA 166  
Db 186 -----VPQVGFQLRHVAHRQVRVVVTGAEAFRHRHVGLLDHR 227  
Qy 167 GVRRLVR-----YVSSVADEGMDCALAAAVYG---GKGGVL---SOLRELKAWFEEQT 214  
Db 228 GGHQVILRAQBDLADLAHQVQVDVAERAAAGNMDVGGQAQVIVQQLRLARQW----- 282  
Qy 215 LFHFYSASILLGYDAAAVAGGCGGVTVKLVDFAHVAEG-----DGVID---HNFLGGL 266  
Db 283 -----VAAAHGADVAVLHQLDVAHLGIGVQRRVVDGEVAAGGGEFLGGL 325  
Qy 267 CSL 269  
Db 326 AAL 328

## RESULT 15

US-09-320-878-1  
; Sequence 1, Application US/09320878A  
; Patent No. 6117659  
; GENERAL INFORMATION:  
; APPLICANT: ASHLEY, Gary  
; APPLICANT: BETLACH, Melanie C.  
; APPLICANT: BETLACH, Mary C.  
; APPLICANT: MCDANIEL, Robert  
; APPLICANT: TANG, Li  
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE  
; FILE REFERENCE: 300622002120  
; CURRENT APPLICATION NUMBER: US/09/320.878A  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908  
; EARLIER FILING DATE: 1998-08-28  
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538  
; EARLIER FILING DATE: 1998-05-06  
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247  
; EARLIER FILING DATE: 1997-04-30  
; EARLIER APPLICATION NUMBER: 60/119,139  
; EARLIER FILING DATE: 1999-02-08  
; EARLIER APPLICATION NUMBER: 60/100,880  
; EARLIER FILING DATE: 1998-09-22  
; EARLIER APPLICATION NUMBER: 60/087,080  
; EARLIER FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 4551  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-320-878-1

Query Match 5.9%; Score 89; DB 3; Length 4551;  
Best Local Similarity 22.1%; Pred. No. 13;  
Matches 62; Conservative 31; Mismatches 106; Indels 82; Gaps 13;  
Qy 4 LHPPEHQVAGHRASASKLGLIDSGSLFYKPLQAGDRGEHVAEYAFSAHA--AVPARI 61  
Db 3732 IHPALLDASLH---AIAVGGVLDEPELVRVFFH-----WSGVTVHAAGAAAARV 3777

Qy 62 RDTFFPRFHGTRLLPTEAQGEPPHVLVLDLL-LAGFQAPCVADIKIGA-----ITWPPSS 116  
Db 3778 R-----LASAGTDAVSLSLDGEGRPLVSVVERLTLRPVTADQAAASRVGGLMHRVAMRPA 3833  
Qy 117 PEYIIAKLAKDRGTT\*SVLLIGFRVSGVRVVGPEGAVRTERPEVKAMDTAGV----- 168  
Db 3834 ----LASSGEQDPHATS-----YGPTAVLGKDELKVAALLESAGVEVGLYPDL 3877  
Qy 169 -----RRVLRVYSSVADEGMDCALAAAVYGGKGVLSOLRELKAWFEEQT 214  
Db 3878 AALSQDVAAGAPAPRTVLAFLPAGPADGGAE-----GVRGTVAITLLELQAWLADE- 3928  
Qy 215 LFHFYSASILL---GYDAAAVAGGCGGVTVKLVDFAHVA 252  
Db 3929 --HLAGTRLLLVTRGAVRDEPGSGADDGG-----EDLSHAA 3962

Search completed: March 24, 2004, 21:00:15  
Job time : 25 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 20:59:15 ; Search time 46 Seconds  
(without alignments)

1626.915 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVAGHRASAK.....IKFVSIVPETHPTQLGPS 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1521	100.0	289	14	US-10-042-894A-8
2	1485	97.6	289	14	US-10-042-894A-8
3	914.5	60.1	227	14	US-10-042-894A-4
4	748.5	49.2	240	14	US-10-042-894A-2
5	742.5	48.8	295	14	US-10-042-894A-14
6	722.5	47.5	310	14	US-10-042-894A-12
7	644.5	42.4	279	12	US-10-424-599-249309
8	644.5	42.4	279	14	US-10-042-894A-10
9	578	38.0	111	14	US-10-042-894A-16
10	577	37.9	111	14	US-10-042-894A-21
11	191	12.6	293	15	US-10-094-749-2660
12	183	12.0	426	9	US-09-731-872-310
13	183	12.0	426	9	US-09-731-872-317
14	183	12.0	426	10	US-09-876-997-310
15	183	12.0	426	10	US-09-876-997-317

16	183	12.0	481	9	US-09-731-872-415	Sequence 415, Appl
17	183	12.0	481	10	US-09-876-997-415	Sequence 415, Appl
18	131	8.6	322	14	US-10-042-894A-23	Sequence 23, Appl
19	120	7.9	33	14	US-10-042-894A-31	Sequence 31, Appl
20	119	7.8	33	14	US-10-042-894A-30	Sequence 30, Appl
21	117	7.7	33	14	US-10-042-894A-33	Sequence 32, Appl
22	116	7.6	33	14	US-10-042-894A-32	Sequence 32, Appl
23	113	7.4	154	9	US-09-864-761-40911	Sequence 40911, A
24	108	7.1	41	14	US-10-042-894A-37	Sequence 37, Appl
25	107	7.0	41	14	US-10-042-894A-36	Sequence 36, Appl
26	105	6.9	316	14	US-10-042-894A-25	Sequence 25, Appl
27	101	6.6	756	14	US-10-156-761-11967	Sequence 11967, A
28	100.5	6.6	660	16	US-10-311-034-26	Sequence 26, Appl
29	100.5	6.6	683	12	US-10-092-900A-342	Sequence 342, Appl
30	98.5	6.5	7349	14	US-10-314-657-46	Sequence 46, Appl
31	98	6.4	41	14	US-10-042-894A-35	Sequence 35, Appl
32	98	6.4	1546	15	US-10-369-493-543	Sequence 543, Appl
33	97	6.4	41	14	US-10-042-894A-34	Sequence 34, Appl
34	93.5	6.1	756	14	US-10-080-170-424	Sequence 424, Appl
35	92.5	6.1	408	14	US-10-156-761-13474	Sequence 13474, A
36	92.5	6.1	486	14	US-10-156-761-9284	Sequence 9284, A
37	92	6.0	308	12	US-10-282-122A-50505	Sequence 50505, A
38	91.5	6.0	320	12	US-10-425-114-47184	Sequence 47184, A
39	91.5	6.0	281	12	US-10-425-114-46878	Sequence 46878, A
40	91.5	6.0	369	12	US-10-425-114-46333	Sequence 46333, A
41	91.5	6.0	375	12	US-10-425-114-42757	Sequence 42757, A
42	91.5	6.0	421	12	US-10-425-114-65310	Sequence 65310, A
43	91.5	6.0	425	12	US-10-425-114-57356	Sequence 57356, A
44	91.5	6.0	428	12	US-10-425-114-46479	Sequence 46479, A
45	91.5	6.0	432	12	US-10-425-114-43357	Sequence 43357, A

## ALIGNMENTS

## RESULT 1

```
US-10-042-894A-8
; Sequence 8, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-042-894A-8
```

Query Match 100.0%; Score 1521; DB 14; Length 289;  
Best Local Similarity 100.0%; Pred. No. 8.2e-150;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSDLPPPHQVAGHRASAKLGPLIDGSGLFYKPLQAGDRGHEHFAFYFAFSAHAAPAR	60
Db	1	MSDLPPPHQVAGHRASAKLGPLIDGSGLFYKPLQAGDRGHEHFAFYFAFSAHAAPAR	60
QY	61	IRDTFFRPHGTRLLPTEAQGPCEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEY	120
Db	61	IRDTFFRPHGTRLLPTEAQGPCEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEY	120
QY	121	IAKCLAKDRGTTSVLLGFRVSGVRVVGPEGAVWRTERPEVKAMDTAGVRRVRLRYSSVA	180

Db 121 IAKLAKDRGTTSVLLGFRVSGVRVVGPEGAVMTERPEVKAMDTAGVRRVLRYYSSVA 180  
QY 181 DEGMDCALAAAVYGGKGVLSQLRELKAWFEEQTLFFHFYSASILLGYDAAVAAGGDDGG 240  
Db 181 DEGMDCALAAAVYGGKGVLSQLRELKAWFEEQTLFFHFYSASILLGYDAAVAAGGDDGG 240  
QY 241 VTKLVDFAHVAEGDGVIDHNFLGGLCSLIKFKVSDIVPETHPTQPLGPS 289  
Db 241 VTKLVDFAHVAEGDGVIDHNFLGGLCSLIKFKVSDIVPETHPTQPLGPS 289

## RESULT 2

US-10-042-894A-6  
; Sequence 6, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase  
; TITLE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 289  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-042-894A-6

Query Match 97.6%; Score 1485; DB 14; Length 289;  
Best Local Similarity 97.9%; Pred. No. 4.6e-146;  
Matches 283; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 MSDLHPPEHQVAGHRASASKLGPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60  
Db 1 MPDLHPPEHQVAGHRASASKLGPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60  
QY 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120  
Db 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFEAPCVADIKIGAITWPPSSPEPY 120  
QY 121 IAKLAKDRGTTSVLLGFRVSGVRVVGPEGAVMTERPEVKAMDTAGVRRVLRYYSSVA 180  
Db 121 IAKLAKDRGTTSVLLGFRVSGVRVVGPEGAVMTERPEVKAMDTAGVRRVLRYYSSVA 180  
QY 181 DEGMDCALAAAVYGGKGVLSQLRELKAWFEEQTLFFHFYSASILLGYDAAVAAGGDDGG 240  
Db 181 DEGMDCALAAAVYGGKGVLSQLRELKAWFEEQTLFFHFYSASILLGYDAAVAAGGDDGG 240  
QY 241 VTKLVDFAHVAEGDGVIDHNFLGGLCSLIKFKVSDIVPETHPTQPLGPS 289  
Db 241 VTKLVDFAHVAEGDGVIDHNFLGGLCSLIKFKVSDIVPETHPTQPLGPS 289

## RESULT 3

US-10-042-894A-4  
; Sequence 4, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase

; TITLE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 227  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-042-894A-4

Query Match 60.1%; Score 914.5; DB 14; Length 227;  
Best Local Similarity 78.5%; Pred. No. 9.4e-87;  
Matches 183; Conservative 8; Mismatches 35; Indels 7; Gaps 2;  
QY 1 MSDLHPPEHQVAGHRASASKLGPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60  
Db 1 MPDLHPPEHQVAGHRASASKLGPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60  
QY 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120  
Db 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFEAPCVADIKIGAITWPPSSPEPY 120  
QY 121 IAKLAKDRGTTSVLLGFRVSGVRVVGPEGAVMTERPEVKAMDTAGVRRVLRYYSSVA 180  
Db 121 IAKLAKDRGTTSVLLGFRVSGVRVVGPEGAVMTERPEVKAMDTAGVRRVLRYYSSVA 180  
QY 181 DEGMDCALAAAVYGGKGVLSQLRELKAWFEEQTLFFHFYSASILLGYDAAVA 232  
Db 180 -----CRRGDGLRARGGGVRRKRWSLVTARAAGVVRGAASVLLGLGVDSSGL 227

## RESULT 4

US-10-042-894A-2  
; Sequence 2, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase  
; TITLE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-042-894A-2

Query Match 49.2%; Score 748.5; DB 14; Length 240;  
Best Local Similarity 64.7%; Pred. No. 2e-69;  
Matches 163; Conservative 10; Mismatches 50; Indels 29; Gaps 6;  
QY 1 MSDLHPPEHQVAGHRASASKLGPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60  
Db 1 MPDLHPPEHQVAGHRASASKLGPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60  
QY 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120  
Db 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFEAPCVADIKIGAITWPPSSPEPY 120  
QY 121 IAKLAKDRGTTSVLLGFRV-----SGVRVV-----GPEGAVMTERPEVKAMDTAGVRRVLRYYSSVA 164

Db 121 IAKLADRGTTSLVGLFRVLRPRRRPRRRVADGAPGCGYGHRRPPRAPALRVILP 180  
Qy 165 TAG-VRRVLRVYSSVADGMDCALAAVYGGKGLVLSQLRELKAWFEQTLFHYASII 223  
Db 181 TRGTARSRRCCTEEKVESCHC--ASSRHGWSRLCS-----TSTRRRFFWAMM 228  
Qy 224 LLGYDAAVAAG 235  
Db 229 LLOSQAQEVGVG 240

## RESULT 5

US-10-042-894A-14  
; Sequence 14, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Caboon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase  
; TITLE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Parthenium argentatum  
US-10-042-894A-14

Query Match 48.8%; Score 742.5; DB 14; Length 295;  
Best Local Similarity 52.1%; Pred. No. 1.1e-68;  
Matches 149; Conservative 45; Mismatches 79; Indels 13; Gaps 5;  
Qy 4 LHPPEHGVAGHRASAKGLPLIDGSLFYKPLQAGDRGEHEVAFYAFSAHAAPARLRD 63  
Db 2 LKAPDHQVAGHEAGLGLKPLIDSGRFYKPLQDNRGESEVAFYESSNNNTPEHR- 60  
Qy 64 TFFPRFHGTRLLPTEAQPGEPPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPIYAK 123  
Db 61 KFFPIYGTIM--KASTGSDHPHVLQDLTSAHVNPSPMDIKIGSRTWAPEAGEAYIAK 118  
Qy 124 CLAKDRGTTSLVGLFRVSGRVVGEA-VWRTERPEVKAMDTAGVRRVLRVYSSVADE 182  
Db 119 CLKXDRSTSLPLGRISGLQYIDDGSGFYKPHRNYMRKTGPADVRLLRKFVSSNPSA 178  
Qy 183 -----GMDCALAAVYGGKGLVLSQLRELKAWFEQTLFHYASIIILGYDAAVAA 234  
Db 179 EMEMRTGLGPCSLASFVYGGNGHIAQLMELKTWFEDQTIYHFACSLFIFEXRLVJK 238  
Qy 235 GGDGGVTVKLVDFAHVAEGDGVLDHNFGLGCLSLIKFVSDIVPPT 280  
Db 239 GA-RSNAEVLIDFAHVTDNGVVDHNFGLGCLSLIKFISDILSET 283

## RESULT 6

US-10-042-894A-12  
; Sequence 12, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Caboon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase

; TITLE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 310  
; TYPE: PRT  
; ORGANISM: Eucalyptus grandis  
US-10-042-894A-12

Query Match 47.5%; Score 722.5; DB 14; Length 310;  
Best Local Similarity 52.8%; Pred. No. 1.5e-66;  
Matches 149; Conservative 38; Mismatches 84; Indels 11; Gaps 6;  
Qy 4 LHPPEHGVAGHRASAKGLPLIDGSLFYKPLQAGDRGEHEVAFYAFSAHAAPARLRD 63  
Db 2 LKVPDHQVAGHRGDDGKGLPLVDDSGRFYKPLQSDHRGDTVEAFYESFYSTELPGLHR- 60  
Qy 64 TFFPRFHGTRLLPTEAQPGE-PPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPIYA 122  
Db 61 KFFPAFHGTXTI--EASDGGSPQPHLVLEDLVSGRTNPPLMDIKTGSRTWYPEASEEYIQ 118  
Qy 123 KCLAKDRGTTSLVGLFRVSGRVV-GEPAVWRTERPEVKAMDTAGVRRVLRVYSSVAD 181  
Db 119 KCLEKDRNSTSVSLGRISGLRVYQNSEAGWQPEKKVYVSFNADGVRSLARKFVSSNLS 178  
Qy 182 EG----MDCALAAVYGGKGLVLSQLRELKAWFEQTLFHYASIIILGYDAAVAAAGD 237  
Db 179 LGPNVDPCLYASKVYCHRGSLAQLQLKEWFEVQVNYHFYSCSLIILYDRESALDG-- 236  
Qy 238 GGGVTVKLVDFAHVAEGDGVLDHNFGLGCLSLIKFVSDIVPE 279  
Db 237 CAHPKVLVDFAHVMGDGHVIDHNFGLGCLSVKIFIRDIAD 278

## RESULT 7

US-10-424-599-249309  
; Sequence 249309, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 249309  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_67156C.1.pap  
US-10-424-599-249309

Query Match 42.4%; Score 644.5; DB 12; Length 279;  
Best Local Similarity 50.7%; Pred. No. 1.7e-58;  
Matches 142; Conservative 35; Mismatches 82; Indels 21; Gaps 9;  
Qy 7 PEHGVAGHRASAKGLPLIDGSLFYKPLQAG----DRGEHEVAFYAFSAHAAPARIR 62  
Db 5 PEHGVAGHKAKDGLGLPLVDDFGFYKPLQNKDDDTGSGTSLSFYSLAA-AAHDYSIR 63  
Qy 63 DTFFPRFHGTRLLPTEAQPGE-PPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPIY 121  
Db 64 -SFFPAFHGTLL--DASDGGSPHPLVLEDLLCGYSKPSVMDVKIGSRTWHLGDSYDI 120

```
QY 122 AKCLAKDRGTTSVLLGFRVSGVRVVGEGAVWTERPEVKAMDTAGVRRVRLRYVSS---178
Db 121 CKCLKKDRSSSLPLGFRISGVK---DSISSWEPTKSLQCLSAHGVALVLNKFFVSSNNI 177
QY 179 -VADEGMDCALAAAVYGGKGVLSQRLKAWFEEOITLFFHYSASILLGVDAAVAAAGD 237
Db 178 NHDDHHPDCAFATEVY---GAVLERQKLDWFEVQTVVHFYSCSVLVVYEKD--LGKGK 232
QY 238 GGGVTVKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSDIV 277
Db 233 ATNPLVKLVDFAHVVDGNGVIDHNFGLGCLSFIFKELKDIL 272

RESULT 8
US-10-042-894A-10
; Sequence 10, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT FILING DATE: 2002-01-09
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Glycine max
US-10-042-894A-10

Query Match 42.4%; Score 644.5; DB 14; Length 279;
Best Local Similarity 50.7%; Pred. No. 1.7e-58;
Matches 142; Conservative 35; Mismatches 82; Indels 21; Gaps 9;

QY 7 PEHQVAGHRASAKGLPLIDSGLFYKPLQAG-----DRGHEVAFYEAFAAHPARIR 62
Db 5 PEHQVAGHKADGGLPLVDVDFGKFKPLQTNKDDTRGSTFSLAA-AAHDYSIR 63
QY 63 DTFPPRHGTRLLPTEAQGE-PPHVLVDLLAGFOAPCVADIKIGAITWPPSSPPYI 121
Db 64 -SFPFAHGTRLL--DASDGSPPHVLVLELLCGYSKPSVMDVKIGRTWHLGDSEDI 120
QY 122 AKCLAKDRGTTSVLLGFRVSGVRVVGEGAVWTERPEVKAMDTAGVRRVRLRYVSS---178
Db 121 CKCLKKDRSSSLPLGFRISGVK---DSISSWEPTKSLQCLSAHGVALVLNKFFVSSNNI 177
QY 179 -VADEGMDCALAAAVYGGKGVLSQRLKAWFEEOITLFFHYSASILLGVDAAVAAAGD 237
Db 178 NHDDHHPDCAFATEVY---GAVLERQKLDWFEVQTVVHFYSCSVLVVYEKD--LGKGK 232
QY 238 GGGVTVKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSDIV 277
Db 233 ATNPLVKLVDFAHVVDGNGVIDHNFGLGCLSFIFKELKDIL 272
```

```
RESULT 9
US-10-042-894A-16
; Sequence 16, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
```

```
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Zea mays
US-10-042-894A-16
```

```
Query Match 38.0%; Score 578; DB 14; Length 111;
Best Local Similarity 97.3%; Pred. No. 4e-52;
Matches 108; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSDLHPPEHQVAGHRASAKGLPLIDSGLFYKPLQAGDRGHEVAFYEAFAAHPARIR 60
Db 1 MPDLHPPEHQVAGHRASAKGLPLIDSGLFYKPLQAGDRGHEVAFYEAFAAHPARIR 60
QY 61 IRDTFFPRFHGTRLLPTEAQGEPPHVLVDLLAGFOAPCVADIKIGAIT 111
Db 61 IRDTFFPRFHGTRLLPTEAQGEPPHVLVDLLAGFOAPCVADIKIGAIT 111
```

```
RESULT 10
US-10-042-894A-21
; Sequence 21, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Zea mays
US-10-042-894A-21
```

```
Query Match 37.9%; Score 577; DB 14; Length 111;
Best Local Similarity 97.3%; Pred. No. 5.1e-52;
Matches 108; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSDLHPPEHQVAGHRASAKGLPLIDSGLFYKPLQAGDRGHEVAFYEAFAAHPARIR 60
Db 1 MPDLHPPEHQVAGHRASAKGLPLIDSGLFYKPLQAGDRGHEVAFYEAFAAHPARIR 60
QY 61 IRDTFFPRFHGTRLLPTEAQGEPPHVLVDLLAGFOAPCVADIKIGAIT 111
Db 61 IRDTFFPRFHGTRLLPTEAQGEPPHVLVDLLAGFOAPCVADIKIGAIT 111
```

```
RESULT 11
US-10-094-749-2660
; Sequence 2660, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
```

```
/ APPLICANT: SUGIYAMA, TOMOYASU
/ APPLICANT: OTSUKI, TETSUJI
/ APPLICANT: WAKAMATSU, AI
/ APPLICANT: SATO, HIROYUKI
/ APPLICANT: ISHII, SHIZUKO
/ APPLICANT: YAMAMOTO, JUN-ICHI
/ APPLICANT: ISONO, YUUKO
/ APPLICANT: HIO, YURI
/ APPLICANT: OTSUKA, KAORU
/ APPLICANT: NAGAI, KEIICHI
/ APPLICANT: IRIE, RYOTARO
/ APPLICANT: TAMECHIKA, ICHIRO
/ APPLICANT: SEKI, NAOHICO
/ APPLICANT: YOSHIKAWA, TSUTOMU
/ APPLICANT: OTSUKA, MOTCYUKI
/ APPLICANT: NAGAHARI, KENJI
/ APPLICANT: MASUHO, YASUHIKO
/ TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
/ FILE REFERENCE: 084335/0160
/ CURRENT APPLICATION NUMBER: US/10/094,749
/ CURRENT FILING DATE: 2002-03-12
/ PRIOR APPLICATION NUMBER: 60/350,435
/ PRIOR FILING DATE: 2002-01-24
/ PRIOR APPLICATION NUMBER: JP 2001-328381
/ PRIOR FILING DATE: 2001-09-14
/ NUMBER OF SEQ ID NOS: 3381
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2660
/ LENGTH: 293
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-094-749-2660

Query Match      12.6%; Score 191; DB 15; Length 293;
Best Local Similarity 26.6%; Pred. No. 3.3e-11;
Matches 68; Conservative 47; Mismatches 83; Indels 58; Gaps 10;

QY 62 RDTFFP-----RFHGTRLLPTEAQGEPPHPLVLDLLAGFOAPCVADIKIGAITWPPS 115
Db 51 RKSFPWGIQCHQAHLTRLC-SEYPENKRRHFLLENVVSQYTHPCVLDLKGTRQHGDD 109
QY 116 SPEPYIAKCLAKDRGTTSVLLGFRVSGVRVGPGEAGVWTERPEVKAMDTAGVRRVRLRY 175
Db 110 ASEKKARHMRKCAQSTACSLGVRCGMQ-----VYQTDKKYFLCKD-----KXY 154
QY 176 VSSVADEGMDCALAAAVYGGK-----GGVLSQLRELKAWPEEQTLFHFYSASILLGYD 228
Db 155 GRKLSVEGPRQADYQFLHNGSHURRLELPILLQLRALLSIIRSQSYRFYSSLLVIYD 214
QY 229 -----AAAVAAGDGGGVY-----VKLVDFAHVA-BG-----DGVIDH 260
Db 215 GQEPPEPAPGSPHPHEAPQAAHGSSPGGLTKVDIRMDIFAHTTYKGWNEHTTYDGP-DP 273
QY 261 NPLGLGLCSLIKFSVDI 276
Db 274 GYIFGLENIIRLIQDI 289

RESULT 12
US-09-731-872-310
/ Sequence 310, Application US/09731872
/ Patent No. US20020102604A1
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, Jean Baptiste
/ APPLICANT: Bougueleret, Lydie
/ APPLICANT: Jobert, Severin
/ TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
/ FILE REFERENCE: 78.US3.REG
/ CURRENT APPLICATION NUMBER: US/09/731,872
/ CURRENT FILING DATE: 2000-12-07
/ PRIOR APPLICATION NUMBER: 60/169,629
/ PRIOR FILING DATE: 1999-12-08
/ PRIOR APPLICATION NUMBER: 60/187,470
/ PRIOR FILING DATE: 1999-12-08
/ NUMBER OF SEQ ID NOS: 426
/ SOFTWARE: Patent.pm
/ SEQ ID NO 317
/ LENGTH: 426
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SIGNAL
/ LOCATION: -28...-1
US-09-731-872-317

Query Match      12.0%; Score 183; DB 9; Length 426;
Best Local Similarity 25.6%; Pred. No. 3.8e-10;
Matches 62; Conservative 42; Mismatches 68; Indels 70; Gaps 7;

QY 88 LVLDLLAGFOAPCVADIKIGAITWPPSPSPYIAKCLAKDRGTTSVLLGFRVSGVRVVG 147
Db 204 ILLENLTSRYEYPCVLDLKMGTROHGDDASEEKAANQIRKCOOSTSAVIGRVCGMQVY- 262
QY 148 PEGAVWRTERPEVKAMDTAGVRRVL---RRYVSSVADEGMDCALAAAVYGGK-----G 197
Db 263 -----QAGSQLMFMNKYHGRKLSMQGFKEALQFFPENGRLRELLG 305
QY 198 GVLSQLRELKAWPEEQTLFHFYSASILLGYD-----AAAVAAG 235
Db 306 PVLKCLTELKAVLERQESYRFYSSLLVIYDGKERPEVVLDSDAETLEDLSSESDESAG 365
QY 236 G-----DGGGVTVKLVDFAH-----VAGSDGVIDHNFGLGSLIKFVSDIYP 278
Db 366 AYAYKPIGASSVDVRMIDFAHTTCRLYGEDTVVHEGQ---DAGYIFGLQSLIDIVTEISE 422
QY 279 ET 280
Db 423 ES 424
```

```
/ PRIOR FILING DATE: 2000-03-06
/ NUMBER OF SEQ ID NOS: 482
/ SOFTWARE: Patent.pm
/ SEQ ID NO 310
/ LENGTH: 426
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SIGNAL
/ LOCATION: -28...-1
US-09-731-872-310

Query Match      12.0%; Score 183; DB 9; Length 426;
Best Local Similarity 25.6%; Pred. No. 3.8e-10;
Matches 62; Conservative 42; Mismatches 68; Indels 70; Gaps 7;

QY 88 LVLDLLAGFOAPCVADIKIGAITWPPSPSPYIAKCLAKDRGTTSVLLGFRVSGVRVVG 147
Db 204 ILLENLTSRYEYPCVLDLKMGTROHGDDASEEKAANQIRKCOOSTSAVIGRVCGMQVY- 262
QY 148 PEGAVWRTERPEVKAMDTAGVRRVL---RRYVSSVADEGMDCALAAAVYGGK-----G 197
Db 263 -----QAGSQLMFMNKYHGRKLSMQGFKEALQFFPENGRLRELLG 305
QY 198 GVLSQLRELKAWPEEQTLFHFYSASILLGYD-----AAAVAAG 235
Db 306 PVLKCLTELKAVLERQESYRFYSSLLVIYDGKERPEVVLDSDAETLEDLSSESDESAG 365
QY 236 G-----DGGGVTVKLVDFAH-----VAGSDGVIDHNFGLGSLIKFVSDIYP 278
Db 366 AYAYKPIGASSVDVRMIDFAHTTCRLYGEDTVVHEGQ---DAGYIFGLQSLIDIVTEISE 422
QY 279 ET 280
Db 423 ES 424

RESULT 13
US-09-731-872-317
/ Sequence 317, Application US/09731872
/ Patent No. US20020102604A1
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, Jean Baptiste
/ APPLICANT: Bougueleret, Lydie
/ APPLICANT: Jobert, Severin
/ TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
/ FILE REFERENCE: 78.US3.REG
/ CURRENT APPLICATION NUMBER: US/09/731,872
/ CURRENT FILING DATE: 2000-12-07
/ PRIOR APPLICATION NUMBER: 60/169,629
/ PRIOR FILING DATE: 1999-12-08
/ PRIOR APPLICATION NUMBER: 60/187,470
/ PRIOR FILING DATE: 2000-03-06
/ NUMBER OF SEQ ID NOS: 482
/ SOFTWARE: Patent.pm
/ SEQ ID NO 317
/ LENGTH: 426
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SIGNAL
/ LOCATION: -28...-1
US-09-731-872-317

Query Match      12.0%; Score 183; DB 9; Length 426;
Best Local Similarity 25.6%; Pred. No. 3.8e-10;
Matches 62; Conservative 42; Mismatches 68; Indels 70; Gaps 7;

QY 88 LVLDLLAGFOAPCVADIKIGAITWPPSPSPYIAKCLAKDRGTTSVLLGFRVSGVRVVG 147
Db 204 ILLENLTSRYEYPCVLDLKMGTROHGDDASEEKAANQIRKCOOSTSAVIGRVCGMQVY- 262
QY 148 PEGAVWRTERPEVKAMDTAGVRRVL---RRYVSSVADEGMDCALAAAVYGGK-----G 197
```

```
Db 263 -----QAGSGQLMFNMKYHGRKLSVQGFKEALFOFFHNGRYLRRELLG 305
Qy 198 GVLSQLRELKAWPEEOTLPHFYSASILLGYD-----AAVAAG 235
Db 306 PVLKLTTELKAVLERQSYRYSSLLVIYDGKERPEVVLDSDAEDLEDSESADESAG 365
Qy 236 G-----DGGGVTVKLVDFAH-----VAEGDGVIDHNFGLGCLSLIKFVSDIVP 278
Db 366 AYAKPIGASSVDVRMIDFAHTTCRLYGEDTVVHEGQ---DAGYIFGLQSLIDIVTEISE 422
Qy 279 ET 280
Db 423 ES 424
```

## RESULT 14

```
US-09-876-997-310
; Sequence 310, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 310
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -28...-1
US-09-876-997-310
```

```
Query Match 12.0%; Score 183; DB 10; Length 426;
Best Local Similarity 25.6%; Pred. No. 3.8e-10;
Matches 62; Conservative 42; Mismatches 68; Indels 70; Gaps 7;

Qy 88 LVLDLLAGFQAPCVADIKIGAITWPPSPPEPIYAKLAKDRGTTSVLLGFRVSGVRVVG 147
Db 204 ILLENLTSRYEVPCLDLKMGTRQHGDDASEEKAANQIRKCOQSTSAVIGRVCGMQVY- 262
Qy 148 PEGAVWTERPEVKAMDTAGVRVL---RRYVSSVADEGMDCALAAAVYGGK-----G 197
Db 263 -----QAGSGQLMFNMKYHGRKLSVQGFKEALFOFFHNGRYLRRELLG 305
Qy 198 GVLSQLRELKAWPEEOTLPHFYSASILLGYD-----AAVAAG 235
Db 306 PVLKLTTELKAVLERQSYRYSSLLVIYDGKERPEVVLDSDAEDLEDSESADESAG 365
Qy 236 G-----DGGGVTVKLVDFAH-----VAEGDGVIDHNFGLGCLSLIKFVSDIVP 278
Db 366 AYAKPIGASSVDVRMIDFAHTTCRLYGEDTVVHEGQ---DAGYIFGLQSLIDIVTEISE 422
Qy 279 ET 280
Db 423 ES 424
```

## RESULT 15

```
US-09-876-997-317
; Sequence 317, Application US/09876997
```

```
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 317
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -28...-1
US-09-876-997-317
```

```
Query Match 12.0%; Score 183; DB 10; Length 426;
Best Local Similarity 25.6%; Pred. No. 3.8e-10;
Matches 62; Conservative 42; Mismatches 68; Indels 70; Gaps 7;

Qy 88 LVLDLLAGFQAPCVADIKIGAITWPPSPPEPIYAKLAKDRGTTSVLLGFRVSGVRVVG 147
Db 204 ILLENLTSRYEVPCLDLKMGTRQHGDDASEEKAANQIRKCOQSTSAVIGRVCGMQVY- 262
Qy 148 PEGAVWTERPEVKAMDTAGVRVL---RRYVSSVADEGMDCALAAAVYGGK-----G 197
Db 263 -----QAGSGQLMFNMKYHGRKLSVQGFKEALFOFFHNGRYLRRELLG 305
Qy 198 GVLSQLRELKAWPEEOTLPHFYSASILLGYD-----AAVAAG 235
Db 306 PVLKLTTELKAVLERQSYRYSSLLVIYDGKERPEVVLDSDAEDLEDSESADESAG 365
Qy 236 G-----DGGGVTVKLVDFAH-----VAEGDGVIDHNFGLGCLSLIKFVSDIVP 278
Db 366 AYAKPIGASSVDVRMIDFAHTTCRLYGEDTVVHEGQ---DAGYIFGLQSLIDIVTEISE 422
Qy 279 ET 280
Db 423 ES 424
```

Search completed: March 24, 2004, 21:04:54  
Job time : 47 secs